DOCUMENT RESUME

ED 056 491 EM 009 347

AUTHOR	Bernstein, B. R.; Gonzalez, B. K.
TITLE	Learning, Retention and Transfer in Military Training.
SPONS AGENCY	Naval Training Device Center, Orlando, Fla.
REPORT NO	TR-69-C-0253-1
PUB DATE	Sep 71
NOTE	95p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Imagery; *Military Training; *Retention; Simulation; Simulators; Skill Development; Task Analysis; Training; *Training Techniques; *Transfer of Training; Visualization

ABSTRACT

The second phase of a project designed to further analyze the factors contributing to the task difficulty in Naval jobs and to determine the optimum training methods for these jobs is reported. A telephone survey of experts attempted to gather data regarding task difficulty. The results of the survey were inconclusive. As a result of the problems experienced in conducting the survey, an overall plan was developed to guide subsequent investigations in which task difficulty is either controlled or manipulated. An experiment was conducted to ascertain whether cartoons, the accompanying text, or a combination of the two were responsible for the facilitating effect of imagery found in an earlier experiment. The results of this experiment failed to validate the findings of the earlier one in that no such facilitating effect was found. A second experiment was therefore conducted, using paper and pencil rather than a hardware simulation of the task. The results of this experiment also showed no evidence that imagery improves performance. A third experiment investigated the influence of training task fidelity of simulation on transfer performance. The data from this study indicated that stimulus and response fidelity has little effect on response time or accuracy. (JY)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED056491

EM009347

Technical Report: NAVTRADEVCEN 69-C-0253-1

LEARNING, RETENTION AND TRANSFER IN MILITARY TRAINING

B. R. Bernstein
B. K. Gonzalez

Honeywell, Inc.
St. Paul, Minnesota 55113

Contract N61339-69-C-0253
NAVTRADEVCEN Task No. 7846-2

September 1971

DOD DISTRIBUTION STATEMENT

Approved for public release and sale; distribution is unlimited.

NAVAL TRAINING DEVICE CENTER
ORLANDO, FLORIDA

TECHNICAL REPORT: NAVTRADEVCEN 69-C-0253-1
LEARNING, RETENTION AND TRANSFER IN MILITARY TRAINING

ABSTRACT

This report summarizes Phase I and reports on Phase II of the L, R, and T program. The purposes of Phase II were to analyze the factors contributing to task difficulty in Naval jobs, and to conduct a further exploration of the influence of imagery, fidelity of simulation, and retention interval on learning, retention, and transfer. A survey of expert opinion regarding task difficulty plus three experimental investigations were conducted during this one-year effort.

The specific purpose of the Task Difficulty Survey was to define those factors which significantly influence task difficulty in the operational Navy setting. Once identified, these factors can be used to manipulate realistically the difficulty of tasks in laboratory investigations. Telephone interviews were conducted with 10 scientists, representing both military and civilian research organizations, to obtain the required information.

The results of this survey were inconclusive. Further, a majority of the difficulty factors mentioned by interviewees were the same ones which have been previously studied in the laboratory. Consequently, the problem became one of suggesting an approach for manipulating such variables. An overall plan is suggested to guide subsequent investigations in which task difficulty is either controlled or manipulated.

Experiments III, IV and V of the current series were conducted during Phase II. Experiments III and IV dealt with imagery and Experiment V with fidelity of simulation.

Experiment III addressed two questions:

1. What is the relative training effectiveness of two forms of imagery; i. e., (a) when new material is related to more familiar material in an analagous fashion and (b) when new material is represented directly?

2. What are the relative contributions of images and text in the presentation of imagery?

This experiment was designed to validate the facilitating influence of imagery found in Experiment II of the series. The findings of Experiment III failed to support the facilitation found in Experiment II. These results are discussed in light of the subject samples and expermental procedures used.

Experiment IV was similar to III, with the exception that subjects were trained with a paper and pencil rather than a hardware simulation of the task. The purposes of this study were to resolve discrepancies in the results between Experiments II and III, and to test the hypothesis that the effectiveness of imagery interacts with verbal skills. The findings were consistent with Experiment III providing no evidence that imagery improves performance. This finding was consistent regardless of verbal skill level. The hypothesis that imagery should be more effective for subjects with lower verbal skills was thus not supported. Those results are discussed in relation to the restricted range of verbal skills of the subjects used.

Experiment V investigated the influence of training task fidelity of simulation on transfer performance. Subjects were trained using one of three levels of fidelity. The data indicated that variations in stimulus and response fidelity had little effect on response time or accuracy. The data are interpreted as supporting the findings of Experiment II; i.e., that relatively high levels of training effectiveness can be obtained for procedural tasks in the absence of high physical fidelity between training and transfer tasks.

Measures of retention examined after two- and seven-day retention intervals (across Experiments I and II, respectively) showed considerable improvement (approximately 20 to 60 percent) in both the speed and accuracy of second-session performance over that of the initial session. However, after a 14-day retention interval (Experiment III) response speed in the second session was no better than first-session speed and for the accuracy measure, performance in the second session was considerably below (approximately 30 percent) the first-session level. The implications of these findings to military training are discussed.

In a final section the work on this program is summarized and recommendations are made for the direction of future research.

GOVERNMENT RIGHTS IN DATA STATEMENT

Reproduction of this publication in whole or in part is permitted for any purpose of the United States Government.

FOREWORD

This report documents a second year's effort to assist the NAVTRADEVCEN in establishing a program of development on learning, retention, and transfer for the purpose of improving the quality and quantity of psychological information pertinent to the design of military training.

THE PRESENT PROGRAM OF RESEARCH

Under the two year contract with Honeywell, Inc., the development program has progressed to a point where, at present, it contains the following main features:

- General information has been gathered and guidelines have been suggested, in a concerted effort made by an outstanding array of experts in the field, especially for the present project.
- Four general types of tasks (procedure following, decision making and problem solving, vigilance, and pattern recognition) have been identified as the most critical to Naval training concerns and are to be subjected to experimental investigation.
- A task designed to represent the procedure-following category has been developed and the study of the learning, retention, and transfer characteristics of "procedure-following" tasks has begun with five experiments. These initial experiments, collectively, incorporate a total of four experimental variables. The first year's research on this project addressed two training variables, viz., "imagery" and "fidelity of simulation." This year's research continued the investigation of these two factors and extended the scope of study to include the additional variables of retention interval and individual differences.
- A telephone survey was conducted in an effort to identify the factors which affect the difficulty level of operational tasks. Such information about the nature of operational tasks will be used, along with the other related information already acquired (i.e., the four general categories of Naval tasks) and additional relevant information that is to be acquired in the future, to relate experimental tasks under investigation to real-world Naval jobs. This aspect of research is considered critical to the objective of placing us in the most favorable position from which to generalize laboratory findings to and facilitate their application in real Naval training situations.

THE PROJECTED PROGRAM OF RESEARCH

Taking the lead provided by the past work on this program, future research will proceed to address task and training variables, some of the highlights of which follow:

TASK VARIABLES. Considered to be of primary importance are the questions of the generalizability and applicability of findings obtained from research on one type of task to various other tasks. These questions should find answers through studies on the influence of the same variables on different task characteristics. In the conduct of such studies, along with the task features already identified under this program (mentioned above), several different schemes for characterizing the features of tasks will be examined for their ability to distinguish among behavioral effects.

TRAINING VARIABLES. Training variables selected for study will be based on certain cognitive notions of emerging interest which emphasize information-processing activities of the performer (of which the imagery variable in the present research is an example). Individual difference factors will be examined where such variables would appear to interact with variables of primary interest. The influence of temporal parameters (e.g., retention interval, duration of practice, distribution of practice) will continue to be explored across experiments, and only in cases of especial interest will such variables be examined within any one experiment. Finally, research on fidelity of simulation will take into account the different possible meanings of the term discussed within (viz., resemblance of training and criterion equipment, resemblance of training and criterion task materials, presence or absence in training of tasks to be performed in transfer).

Arthur S. Blaiwes
ARTHUR S. BLAIWES
Human Factors Laboratory

TABLE OF CONTENTS

Section Page

I INTRODUCTION 1
  Background 1
    Experiment I 2
    Experiment II 3
    Phase I Conclusions 7
  Statement of the Problem 7
    Identification of Task Difficulty Factors 7
    L, R, and T Experiment III 8
    L, R, and T Experiment IV 9
    L, R, and T Experiment V 9

II TECHNICAL APPROACH - A SURVEY TO IDENTIFY FACTORS INFLUENCING TASK DIFFICULTY 11
  Method 11
  Problems Encountered 12
  Results and Discussion 12
  Recommended Approach 15
  Conclusions 16

III TECHNICAL APPROACH - EXPERIMENT III 17
  Method 17
    Apparatus 17
    Subjects 20
    Procedure 20
    Scoring 26
  Results 26
  Discussion and Conclusions 29

IV TECHNICAL APPROACH - EXPERIMENT IV 32
  Method 32
    Apparatus 32
    Subjects 32
    Procedure 32
  Results 33
  Discussion and Conclusions 33

V TECHNICAL APPROACH - EXPERIMENT V 38
  Method 38
    Apparatus 38

Subjects 41
Procedure 41
Experimental Design 41
Results 42
Discussion and Conclusions 42
Integrative Review of Imagery and Fidelity 47

VI TECHNICAL APPROACH - RETENTION EFFECTS 53
Overall Savings Versus First Trial Savings 55

VII SUMMARY AND RECOMMENDATIONS 56
Imagery 56
Fidelity of Simulation 57
Task Difficulty 58

REFERENCES 60

APPENDIX A SUMMARY OF INTERVIEW ON TASK DIFFICULTY 65

APPENDIX B DIRECT IMAGERY (L, R, AND T) 68

LIST OF ILLUSTRATIONS

Figure		Page
1	Subject's Console	18
2	Experimenter's Console	19
3	L, R, and T Console	19
4	Percent Accuracy for Acquisition and Transfer Sessions	27
5	Mean Response Time for Acquisition and Transfer Session	27
6	Mean Response Time for Acquisition and Transfer Sessions of Experiments II and III	28
7	Mean Time to Complete Subroutines (Verbal Skill Groups)	34
8	Mean Time to Complete Subroutines (Imagery Levels)	34
9	Mean Percent Accuracy	35
10	Percent Accuracy	35
11	High-Fidelity Panel	38
12	Medium-Fidelity Panel	39
13	Low-Fidelity Panel	40
14	Mean Response Time as a Function of Stimulus Fidelity	43
15	Mean Response Time as a Function of Response Fidelity	43
16	Accuracy as a Function of Stimulus Fidelity	44
17	Accuracy as a Function of Response Fidelity	44
18	First Block and Average Savings Based on Response Time Scores, as a Function of Retention Interval	54
19	First Block and Average Savings Based on Response Accuracy Scores, as a Function of Retention Interval	54

LIST OF TABLES

Table		Page
1	Factors Suggested to Influence Task Difficulty	13
2	Types of Problems Used in L, R, and T Experiments I and II	21
3	Experimental Design - L, R, and T Experiment III	25
4	Analysis of Variance Summary - Response Times During Training Session	28
5	Analysis of Variance Summary - Response Times During Transfer Session	30
6	Analysis of Variance Summary - Response Accuracy During Training Session	30
7	Analysis of Variance Summary - Response Accuracy During Transfer Session	30
8	Analysis of Variance Summary - Response Time During Training	36
9	Analysis of Variance Summary - Response Accuracy During Training	36
10	Experimental Design - L, R, and T - Experiment V	41
11	Analysis of Variance Summary - Response Time During Training	45
12	Analysis of Variance Summary - Arcsin Transformed Percent Accuracy Scores During Training	45
13	Analysis of Variance Summary - Response Time During Transfer	46
14	Analysis of Variance Summary - Arcsin Transformed Percent Accuracy Scores During Transfer	46

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the assistance of Miss Jenny Crowe, Research Assistant at Honeywell, in collecting and analyzing data.

They also wish to thank P. W. Fox, Ph.D., Psychological Consultant for Honeywell at the University of Minnesota, for his valuable inputs.

SECTION I

INTRODUCTION

BACKGROUND

This report summarizes the results of the second year of work in the study of learning, retention, and transfer (L, R, and T), applicable to the design and utilization of Naval training systems. During the initial phases of the program, Honeywell contracted to provide assistance to the NAVTRADEVCEN in the establishment of a long-term in-house research activity. Due to the wide scope of L, R, and T research and the limited resources available, considerable planning has been required to define an effort of manageable proportions.

The formulation of a research plan was accomplished as the result of the following activities (Bernstein and Gonzalez, 1969):

- A brief review of relevant literature
- Selection of an appropriate task taxonomy
- A survey of Navy jobs to determine critical tasks from the standpoint of training
- Organization and management of technical meetings with consultants
- A conceptual design for a general-purpose L, R, and T apparatus

As development of this plan took the greater portion of the first year, only two experimental studies were conducted. These studies were designed to try out the newly developed set of program requirements. The second year's efforts emphasized laboratory investigations to provide a data base upon which subsequent research can build. In addition, an analysis of factors contributing to task difficulty in operational jobs was performed. This analysis logically followed the survey of Naval jobs conducted during Phase I.

A basic problem in conducting L, R, and T research is choosing from a long list of possible variables which can be studied. The selection of problems for study should be based on a priority system. The program plan was formulated with an aim toward establishing priorities for research topics.

A variable of central importance is the taxonomic category of the skill for which L, R, and T data are to be collected. The number of possible behavioral categories varies with the taxonomy used. Part of the planning phase involved the selection of a taxonomy for specifying tasks. That of Willis and Peterson (1961) was chosen because it was developed specifically for Naval training applications. Willis and Peterson define tasks at three levels of specificity. The broadest level contains three categories. At the second level, these three are broken down to a list of six. Nineteen task/behavior categories are given in the level III list.

On the basis of Willis and Peterson's descriptions and examples of level III categories, an analysis of Naval jobs was performed to select high-priority tasks for initial investigation. This analysis suggested that procedural skills were required for the performance of virtually all critical Navy tasks. The decision was made to conduct initial investigations on the L, R, and T of procedure-following tasks.

An apparatus was designed and fabricated to serve as an experimental procedural skills task. The device resembles, in a number of ways, military electronics equipment. Operation of this equipment requires subjects to learn and apply facts, principles, and procedures in a synthetic communications task. Performance is measured in terms of response latency and accuracy. Based on an analysis of task sequences, an error analysis, and interviews with subjects, scores reflected a minimum of skill components other than procedure-following.

Experiment I

The major purpose of this pilot study was to establish baseline data on the experimental task for future investigations. More specifically, data from this experiment were needed for the following purposes.

- The identification of methodological problems.
- Determination of acquisition and transfer performance functions, as measured by time and accuracy scores, and the correlation between these measures.
- Because a transfer of training design was to be used throughout the program, two sets of equivalent problems were required. Experiment I was designed to compare performance on these two sets of problems. Experiment I also allowed blocks of problems to be equated in difficulty.
- Information was needed about how subjects organized information necessary to performing the task. During Experiment I, post-experimental verbal reports were taken with the data used to make final improvements in the instruction manual.

These data were also used as the basis for forming imagery sequences, in an effort to improve training effectiveness. Imagery will be discussed in more detail later in this section.

Time and accuracy scores during acquisition conformed to the classical negatively accelerated monotonic function, characteristic of more traditional learning experiments. Improvement over problem blocks was associated more with time than with accuracy scores. Moreover, the correlation between the two measures was not statistically significant, suggesting independence between measures. These data were interpreted as evidence that subjects tended to work slowly until they became more confident of the correct procedures.

An absence of significant differences between performance on the two sets was found. Thus, two equivalent problem sets were available for use with Woodworth and Schlosberg's (1954) Plan 3 or Plan 5 for transfer of training experiments.

An examination of subjects' post-experimental verbal reports directed attention to a familiar psychological principle. Several subjects stressed the need to understand basic functional relationships in the system. These subjects indicated that the use of correct procedures was easy if one fully understood the nature of the communications network. It was interesting that the subjects who volunteered this information performed at generally higher levels than did the other subjects.

Another interesting finding was the comparison between performance during the first and second sessions. This investigation showed no evidence of forgetting over the 24-hour retention interval between sessions. The absence of a retention loss on this task contrasts strongly with the findings of more typical memory experiments using verbal materials. The difference may be associated with the meaningful organization of verbal associations found in the present task. Such organization tends to be absent in rote list learning tasks, unless it is imposed mnemonically by the subject. When such a mnemonic system is used in list learning, forgetting is drastically reduced (e.g., Senter and Hauder, 1968).

An appeal to inherent organization in the to-be-learned material of this procedural task is consistent with the earlier discussion of subjects' verbal reports. That is, subjects were unlikely to forget the nature of the system in one day. Thus, they could use their understanding of functional system relationships as pegs for recalling specific details of operation.

Experiment II

Two variables were selected for initial investigation in Experiment II, and to a further extent, during the second year of the program. The decision to study imagery and fidelity of simulation was primarily based on discussions held at the L, R, and T Technical Meetings.*

---

*The subject matter of these meetings is provided in Appendix B of the Phase I final report (Bernstein and Gonzalez, 1969).

Fidelity of Simulation -- This term refers to the similarity between a training device and its operational counterpart. Similarity can vary both in terms of the stimuli and the responses involved in equipment operation. Thus, fidelity of simulation can be viewed within a transfer of training paradigm.

The designers of training devices tend to assume that training effectiveness varies as a function of the degree to which the trainer physically simulates the operational equipment. This assumption is consistent with the classical model of transfer (Osgood, 1949) to the extent that positive transfer is highest when both the stimuli and responses of two tasks are maximally similar.

One may question, however, the applicability of Osgood's model to the design of even moderately complex training systems. Two major differences exist between Naval tasks and the kinds of tasks used to test the Osgood paradigm. First, the typical transfer of training experiment involves a simple relationship between the stimuli and responses of the acquisition and transfer tasks. Second, as was mentioned earlier, laboratory learning tasks tend to differ fundamentally from real-world tasks in forms of inherent organization.

To illustrate both points, let us consider a typical transfer experiment. The majority of transfer investigations have been conducted using a paired-associate (PA) task. The reason for this is that easily identifiable stimulus and response terms are present in a PA list. Maximum positive transfer is obtained when subjects are trained to provide the correct response term (RT) for each stimulus term (ST), and the same list is carried over to the transfer situation. Maximum negative transfer takes place when both the ST's and RT's are carried over, but the associations are reconstructed on a random basis. In the latter situation, the subject must both unlearn the original associations and acquire the new ones. Merely carrying over the ST's and replacing the RT's with an entirely new set typically fails to produce negative transfer, and often actually results in a small amount of positive transfer (Battig, 1966). According to Battig's (1968) analysis of the specific components of learning in a PA task, it is not surprising that positive transfer is obtained in this last situation. That is, many components of learning are carried over from acquisition to transfer.

The difficulty in analyzing even rather simple equipment operating tasks into a list of PA's is evident. Operational procedural tasks usually involve sequential dependencies not present in an orthodox PA list. The task used in the present program has a logical structure with responses contingent on stimuli and other responses.*

---

*See Appendix D of the Phase I report (Bernstein and Gonzalez, 1969).

The presence of a logical structure in this procedural task produces an apparent paradox. In a sense there is a complexity factor associated with the branching of contingent responses. However, the fact that the contingencies have an inherent logical basis tends to simplify operation. This point relates to the earlier discussion about Experiment I, where we point out the importance of learning functional system relationships, compared with rote learning of procedures. Without a well formulated mnemonic system, the subject in a PA experiment has no choice but to learn by a rote method. This discussion leads to the hypothesis that transfer, in a realistic procedural task, does not critically depend on relationships between specific stimuli and specific responses. Rather, it depends more on the extent to which knowledge of system structure and functional relationships applies to the operational task. If the training system does an effective job of teaching underlying system structure, trainees may be able to tolerate relatively low levels of physical fidelity.

Wittrock (1968) suggests that transfer can be achieved through mediated generalization. He reasons that man can generalize and discriminate among different situations by abstracting critical features from those situations and by generating verbal descriptors. These verbal descriptors then mediate the physical response. One would thus hypothesize that even if the stimuli and responses of the task were physically different, but the trainee was able to abstract the critical elements, transfer would remain at a high level.

Experiment II was designed in part as a preliminary test of this hypothesis. One group of subjects received training on the operational equipment. A second group was trained on a paper and pencil representation of the equipment. The front panel drawing was highly accurate, but was rendered in black and white. Color coding was used to denote priority information in the actual device, while priority data in the low fidelity trainer was conveyed by letter designations. Both groups received 30 acquisition problems using either the high or low fidelity device.

The primary difference between the two fidelity conditions was in the way subjects responded. Low fidelity subjects could not actually operate controls, but instead, wrote in their responses on an answer sheet. For example, instead of turning a rotary switch, the subject wrote the number of the switch position; instead of pushing switches, the subject made a check mark over the proper switch, etc.

After a one-week retention interval, all subjects received a second set of 30 problems on the operational equipment. Low-fidelity subjects exhibited significantly longer response times on the first set of transfer problems, but thereafter did not differ from the high-fidelity group. The two groups were equivalent, throughout, in terms of accuracy.

These results were consistent with the experimental hypothesis and the findings of other investigations (e.g., Grimsley, 1969; Mahler and Bennet, 1950; Newton, 1959; Wilcoxin, Davy and Webster, 1954); i.e., relatively low fidelity representations of the task produced satisfactory training effectiveness. While each of the other studies involved actual military tasks, containing an uncontrolled combination of skill components, the present study concerned itself solely with procedural learning and transfer.

Imagery -- Imagery, as related to retention, denotes the use of visual mental representations of relatively concrete objects as mediators for storage. The term has recently been used in connection with mnemonic systems for improving memory. Mnemonists are able to recall very long lists of unrelated items on a one-trial learning basis. They accomplish this by associating the to-be-learned items with a previously overlearned set of word pegs, each having a basic logical relationship with every other one (Yates, 1966). The association of the new items with the pegs is thought to be enhanced by the creation of usually bizarre images of a relationship.

Recently psychologists have investigated the use of imagery under controlled laboratory conditions. Senter and Hauder (1968) found that subjects who were trained to use the above-described mnemonic system exhibited markedly superior acquisition and retention performance over a control group which was not given imagery training.

Paivio and his colleagues have conducted a number of experiments on the image-arousing potential (*I*) of word pegs (e.g., Paivio, 1963; Paivio, Yuille, and Madigan, 1968; Paivio, Yuille and Smythe, 1966). Collectively, these experiments provide convincing evidence that *I* is an important determinant of speed of acquisition when subjects are instructed to use imagery. Moreover, *I* acts in much the same way as meaningfulness, but relates more to the concreteness of the word pegs.

The post-experimental verbal reports of subjects from Experiment I combined with the concept of imagery suggested an experimental technique for improving training effectiveness. The idea was to provide subjects with a more concrete means of acquiring and retaining knowledge of system structure. The technique was based on the fact that imagery involves two principal features: 1) a set of very familiar items acting as pegs for the mediation of the new material, and 2) highly vivid visual representations of a relationship between the pegs and the to-be-acquired items.

The present application of imagery differed from the usual mnemonic methods because the material to be learned was not simply a list of items. Rather, the aim was to present a clear description of how the communication system was structured. We reasoned that such information could be made more salient if an analogy could be drawn between this system and a system with which subjects had basic familiarity.

Subjects in the imagery group were given a five-minute presentation of cartoon slides showing situations occurring in the course of commercial air travel which were analogous to each of the communication routines. The text of the commentary accompanying each slide was contained in the manual and read verbatum to the subjects. Control group subjects were given the time used for imagery presentation (5 minutes) to study the manual.

The imagery group exhibited superior response times in the acquisition and transfer sessions. The time saved by imagery subjects averaged about 25 percent. Imagery produced no significant differences in either session when performance was measured in terms of percent accuracy.

Phase I Conclusions

The results of the two initial experiments suggested the following four conclusions:

- Modifications to training methods, based upon feedback from trainees, can increase training effectiveness.
- Providing subjects with fundamental knowledge of basic system structure and their role within that structure seems to be more critical to performance than simply teaching facts, principles, and procedures.
- Imagery may be a valuable technique for providing knowledge of system structure.
- Apparently, training effectiveness is not reduced by low-fidelity training on the procedural task used in the experiments.

STATEMENT OF THE PROBLEM

Phase II of the L, R, and T program was concerned with: (1) an analysis of factors which contribute to task difficulty in Naval jobs, and (2) a further examination of imagery and fidelity of simulation. This one-year effort involved three experiments and a survey of expert opinion regarding task difficulty. L, R, and T Experiments III and IV dealt with imagery, and Experiment V with fidelity of simulation.*

Identification of Task Difficulty Factors

Consistent with the pragmatic approach adopted during the first year of this program, efforts have been made to select research topics with maximum relevance to operational training problems. Although the earlier survey of

---

*What is now called Experiment IV was actually the fifth experiment, chronologically, but because it involved the study of imagery, it is presented directly after Experiment III.

Navy jobs produced a list of high-priority tasks, the survey did not identify factors influencing the difficulty of performing these tasks.

The specific purpose of the Task Difficulty Survey was to define those factors which influence task difficulty in the operational setting. Once defined, these factors can be used to manipulate task difficulty in the laboratory setting. This approach to the selection and use of variables for laboratory studies stresses maximum relevance and transferability of laboratory findings to the solution of real-world problems.

L, R, and T Experiment III

The purpose of this investigation was to explore further the role played by imagery in the acquisition and transfer of procedural skills. Experiment II suggested that the principle of imagery has considerable potential to enhance procedural learning; i.e., learning is facilitated when subjects can relate to-be-learned material to previously familiar information through visualization. In the earlier study, the familiar material acting as a peg was the commercial air-travel system. Subjects who were shown cartoons depicting this system and were told how various travel situations were similar to analogous communication procedures operated the synthetic communication device more rapidly than did an appropriate control group. This application contained both fundamental ingredients of the imagery principle, viz., the relations between unlearned and overlearned information and the provision of visual images. Although both ingredients seem to be necessary in the more traditional use of imagery as a mnemonic device in verbal learning, one may question their criticality in a procedural skills context.

The purpose of Experiment III was to provide information on this issue. More specifically, two questions were addressed by the present investigation, viz:

- What is the relative training effectiveness of two forms of imagery; i.e., (1) when new material is related to some familiar material in an analogous fashion, and (2) when new material is represented directly.

- What are the relative contributions of images and text in the presentation of imagery?

Experiment III was designed to ascertain whether the cartoons, the accompanying text, or a combination of the two were responsible for the facilitating effect of imagery found in Experiment II. Moreover, a direct comparison was made between imagery directly related to the communication system, and imagery describing an analogous situation. The comparison tests the value of pegs, which historically have been overlearned material. This is in contrast to providing word pictures of the to-be-learned material, which may themselves mediate the element of learning.

L, R, and T Experiment IV

Training subjects in the operation of the L, R, and T communication device or training operators of Naval systems in actual procedural tasks requires relatively large amounts of verbal instruction. Even moderately complex procedural tasks involve a host of contingent actions. The presentation of these contingencies and the correct responses for each, may result in many pages of text or lengthy direct discourse. One would thus suspect that a trainee's ability to acquire such information would be heavily influenced by his verbal skills. In other words, a trainee's ability to follow instructions is limited by his ability to understand the discourse as it is presented.

It was mentioned earlier that imagery may be effective because it tends to increase the concreteness of the to-be-acquired material. Thus, it would seem that the imagery would be particularly effective for trainees who have difficulty comprehending verbally presented instructions.

The purpose of Experiment IV was to test the hypothesis that there is an interaction between verbal skill level and the effectiveness of imagery. Such an experiment requires measurement of this skill and the blocking of subjects according to their test scores.

The most common tests of intelligence contain subtests for verbal skills. Verbal intelligence correlates highly with academic achievement: a learning experience which requires comprehension based on verbally presented material. A case can thus be made for assuming that acquisition rate on a procedural skills task, similar to that of the present experiment, would correlate with verbal intelligence. This reasoning led to the use of a verbal IQ test. The Ammons Quick Test (Ammons and Ammons, 1962) was selected as a basis for blocking verbal abilities. This test was chosen primarily because of the speed with which it can be administered and the fact that it correlates highly with more traditional verbal abilities tests.

L, R, and T Experiment V

The major difference between high and low fidelity of simulation in Experiment II was in the way subjects responded to communication problems. The low-fidelity condition involved a rather accurate stimulus simulation of the criterion equipment. The absence of degradation in training effectiveness with the low fidelity training conditions suggests that practicing the identical physical response is not critical to the transfer of procedural skills.

The next logical question involves similarity of the equipment used in the training and transfer situations. One might ask how much difference in appearance subjects can tolerate between training equipment and operational hardware.

This question was tested experimentally in the present investigations. However, the changes in stimulus conditions were made with a very fundamental restriction. That is, the basic functional relationships in the communications systems always remained constant from training to transfer.

These kinds of variation are analogous to situations in which the front panels of electronics systems are periodically redesigned, but the electronics behind the panels remain essentially undisturbed. Do such design changes cause training equipment, based on older operational equipment, to become obsolete?

Based on Wittrock's (1968) concept of mediated generalization, a prediction similar to that made in Experiment II seems to be justified. Subjects should not be affected by differences in physical stimuli, just as they were not substantially affected by response differences. Again, because knowledge of system structure mediates the acquisition of detailed procedures, subjects ought to be able to tolerate changes in physical appearance when the fundamental system is not changed.

Another aspect of the present investigation again relates to the way subjects respond during the execution of training problems. If verbal mediation is the critical determinant of transfer, it may not be necessary for trainees to operate the equipment when learning. That is, it may be sufficient for trainees to merely verbalize what responses they would make if they were operating the equipment. This hypothesis was tested by comparing two groups of subjects. Subjects in the experimental group told the experimenter what responses they would make and the control group wrote in their responses.

SECTION II

TECHNICAL APPROACH - A SURVEY TO IDENTIFY FACTORS INFLUENCING TASK DIFFICULTY

METHOD

Honeywell originally proposed to obtain task difficulty information by visiting and interviewing personnel in the field. Because of the anticipated scope of such an effort, NAVTRADEVCEN suggested that this information might be obtained from individuals who have previously conducted studies in the operational environment. Therefore, it was agreed that Honeywell would conduct a telephone survey of selected scientists to gather the required information, if possible, without visiting operational systems.

The persons contacted during this survey were:

H. Ammerman - HumRRO, Ft, Bliss, Texas
G. Briggs - Ohio State University, Columbus, Ohio
G. Jeantheau - Dunlap and Assoc., Darien, Conn.
R. Mackie - Human Factors Research, Santa Barbara, Calif.
T. Mara - General Dynamics, Electric Boat Div., New London, Conn.
W. Matheny - Life Sciences Inc., Ft. Worth, Texas
E. Miller - HumRRO, Ft. Bliss, Texas
L. Schrenk - Honeywell Research, St. Paul, Minn.
A. Siegel - Applied Psychological Services, Wayne, Pa.
E. Yaeger - Honeywell Marine Systems Center, West Covina, Calif.

The discussion of task difficulty centered on each of the interviewee's responses to the following two questions:

1) Do you feel it is of value to visit the operational setting to obtain information about task difficulty?

2) On the basis of your experience (i.e., familiarity with military jobs and experimental investigations), what factors do you feel influence task difficulty in the operational setting?

The analysis of interviewee responses involved: (1) determination of general factors contributing to task difficulty; (2) identification of those specific tasks which respondents felt to be highly difficult; and (3) identification of the major factors producing difficulty for these specific tasks.

21

PROBLEMS ENCOUNTERED

Two major problems restrict the usefulness of information obtained. First, the interviewees responded in idiosyncratic terms to describe task difficulty factors. The different interviewees used different terminology to describe similar factors, for example, the "number of different things to be done" may mean the same as "system load/overload". However, without operational definitions for these terms, it was impossible to state unequivocally what was meant.

A second problem concerned the experimental backgrounds of interviewees. Although each person contacted is a recognized human factors specialist, the background experience of each of these men has by no means covered all tasks. Opinions regarding task difficulty were, therefore, based on different types of observed behavior. This again limits the extent to which the views of one interviewee can be compared with another.

RESULTS AND DISCUSSION

Without exception, interviewees responded affirmatively to Question No. 1. Several respondents cautioned, however, that visits to the operational setting must be preceded by adequate preparation, i.e., complete system familiarity and well formulated hypotheses. Concern was also expressed about the ability of observers to establish and maintain good rapport with operational personnel. This factor is considered critical if meaningful data is to be gathered. The consensus was that conduct of effective field investigations tends to be more of an art than a science, since rigorous control procedures seldom are possible to implement. Lack of control is based on the fact that this type of study must be conducted, for the most part, on a noninterference basis. One must therefore accept data under varying sets of circumstances.

To place responses to Question No. 2 in clearer perspective, a matrix of responses was developed (Table 1). This table summarizes the factors suggested by interviewees as producing task difficulty. The individual interviewee responses are elaborated in Appendix A. An examination of the appendix indicates the degree of arbitrariness that was necessary in developing this table; it is apparent that other categories could have been used. In addition to the lack of agreement between investigators, Table 1 and Appendix A indicate that the majority of the factors mentioned are those which have been frequently studied in the laboratory. If a competent experimental psychologist would be asked to list task difficulty factors, it is likely that the list would include task organization, complexity, ambiguity, memory, etc. Therefore, based on this small sample, going to the operational environment to catalog variables which affect task difficulty appears to be inefficient.

Table 1. Factors Suggested to Influence Task Difficulty

	Number of Decision Options	Number of Different Things to be Done	Amount of Processing Required	Task Organization	System Load Overload	Ambiguity	Amount of Information Available	Amount of Storage Available	Stress	Adaptation Requirements	Coordination	Complexity	S - - N Ratio
Ammerman	X	X						X	X		X	X	
Briggs		X	X	X								X	
Jeantheau		X			X				X				
Mackie													X
Mara						X	X	X					
Matheny					X					X			
Miller		X						X	X		X	X	
Schrenk						X	X					X	
Seigel											X		
Yaeger												X	

- Performance criteria
- Operater characteristics
- Task complexity
- Training characteristics

These task difficulty factors identified above can all be subsumed under the variable of task complexity. Difficulty, however, is considerably broader than this, and varies, regardless of task, as a function of the following variables:

- Performance criteria
- Operator characteristics
- Task complexity
- Training characteristics

It is felt therefore, that difficulty is relative, in that no task is difficult in and of itself. By way of example, Folley and Munge · (1961) have shown that performance can be dramatically influenced by varying the "degree of proceduralization" of a task. That is, within a single task there can be varying levels of difficulty in performance.

There is a problem of definition and scaling which relates to all of the above difficulty factors. Typically, operational definitions and unique measurement schemes have been prepared each time a task difficulty variable is used in a study. Obviously one can continue to define terms and measures to fit each specific investigation. However, due to their lack of generalizability, such procedures do not provide a workable solution to our problem.

The most desirable solution involves the development and validation of general difficulty factor definitions and measures. (Despite the fact that much study has been devoted to the area of psychometrics , much remains to be done.) Moreover, we consider this to be a task which is far beyond the scope and resources of the present contract.

In general, task performance can be improved by any of the following procedures:

- Reducing the criteria for satisfactory performance
- Improving the selection procedures, thereby increasing the input skill level
- Reducing task complexity by redesign or incorporation of performance aids
- Improving the effectiveness of training techniques or increasing the length of training.

Further, levels for each of these factors can be defined thus establishing a range of performance comparable to that which could be expected under operational circumstances. Considerably more information is still required about the effective levels and possible interactions among these factors. Nevertheless, translatable research can be accomplished if the experimental situation is carefully patterned after the operational environment.

RECOMMENDED APPROACH

The experimental approach which we propose for implementation is based on the notion that task difficulty should not be treated as a unitary independent variable. Rather, the term should be used as an intervening variable and studied as it is affected by the manipulation of performance criteria, operator characteristics, task complexity, and type of training.

The primary advantage of this approach is that laboratory research data can be applied to the design of training for Naval jobs by experimentally manipulating factors which directly bear upon successful performance. That is, tradeoffs can be identified to weigh training variables against each of the other three factors. For example, the performance of a particular task, by a certain class of personnel, has been shown to be unacceptable, regardless of the quality of instruction. It would make little sense to continually strive to develop new training procedures when personnel selection or task design is the limiting factor in job performance. Laboratory data can provide estimates of performance at various combinations of levels of each of the four factors which influence task difficulty.

Implementation of this approach requires the development of an overall study plan for investigating task difficulty. This study plan must identify specific levels of the four major variables for each experiment. Specification of these levels should be aided by knowledge of the operational setting. That is: criteria, task complexity, personnel selection, and training procedures can be modeled after those used in the Navy. In so doing it should be possible to establish a basis of comparison for evaluating experimental procedures associated with each factor.

Although the approach proposed here is seen as having a number of distinct advantages, it is not considered to be a panacea. Perhaps the major advantage to our approach is the development of a theoretical and systematic framework for investigating performance difficulty. Of course, this approach to task difficulty requires additional time and resources to manipulate the larger number of variables systematically. However, the proposed systems approach should have fewer undesirable influences on the overall program than will the alternative of establishing operational definitions and unique measurements for difficulty factors each time a study is conducted.

CONCLUSIONS

1) The current survey failed to provide desired information concerning factors which influence task difficulty.

2) The specific difficulty factors mentioned by interviewees did not differ significantly from those which might be considered if a researcher were to simply make a list of such factors.

3) There is a lack of generalizable definitions and measurement schemes for the various factors which influence task difficulty, i. e., task complexity, performance criteria, operator skill level, and training methods.

SECTION III

TECHNICAL APPROACH - EXPERIMENT III

METHOD

Apparatus

The experimental task in Experiment III required the subject to operate a synthetic communication system. Two units of equipment constituted the apparatus; the subject's console (Figure 1) and the experimenter's console (Figure 2). Figure 3 shows the complete laboratory facility.

Equipment operated by the subject included a panel of indicators and switches (communications control console) and an information display. The communications control console (CCC), located directly in front of the seated subject, housed all displays and controls necessary for operating the system. The information display, which contained all the facts, principles, and procedures was a Kodak random-access carousel slide projector and a 5 x 7-inch back-illuminated Polacote screen. Subjects accessed 35 mm slides with a selector located at the bottom left of the CCC. An index slide provided information about the location of all other slides.

The experimenter's control panel was connected electrically to the CCC. The experimenter's unit served two functions: It permitted selection of the communications problems to the subject, and it provided to the experimenter information concerning the subjects' responses. A Hunter timer, positioned at the left of the experimenter's panel, was used to record the lapsed time to 0.1 second, from the presentation of a problem until its completion.

The CCC was arranged in three sections. At the left of the panel were three displays (IEE readouts) which presented input information. One display showed the designating number of the initiator of the communication (sender). Below the sender's number was the receiver's number. At the bottom left of the CCC was a readout which contained the priority of the requested transmission.

The central part of the CCC contained 18 microswitch trans-illuminated pushbuttons. The top three switchlights were used to select a frequency for communication between the sender and the CCC. Below the frequency selectors was a 3 x 5 matrix of switchlights which served the following functions:

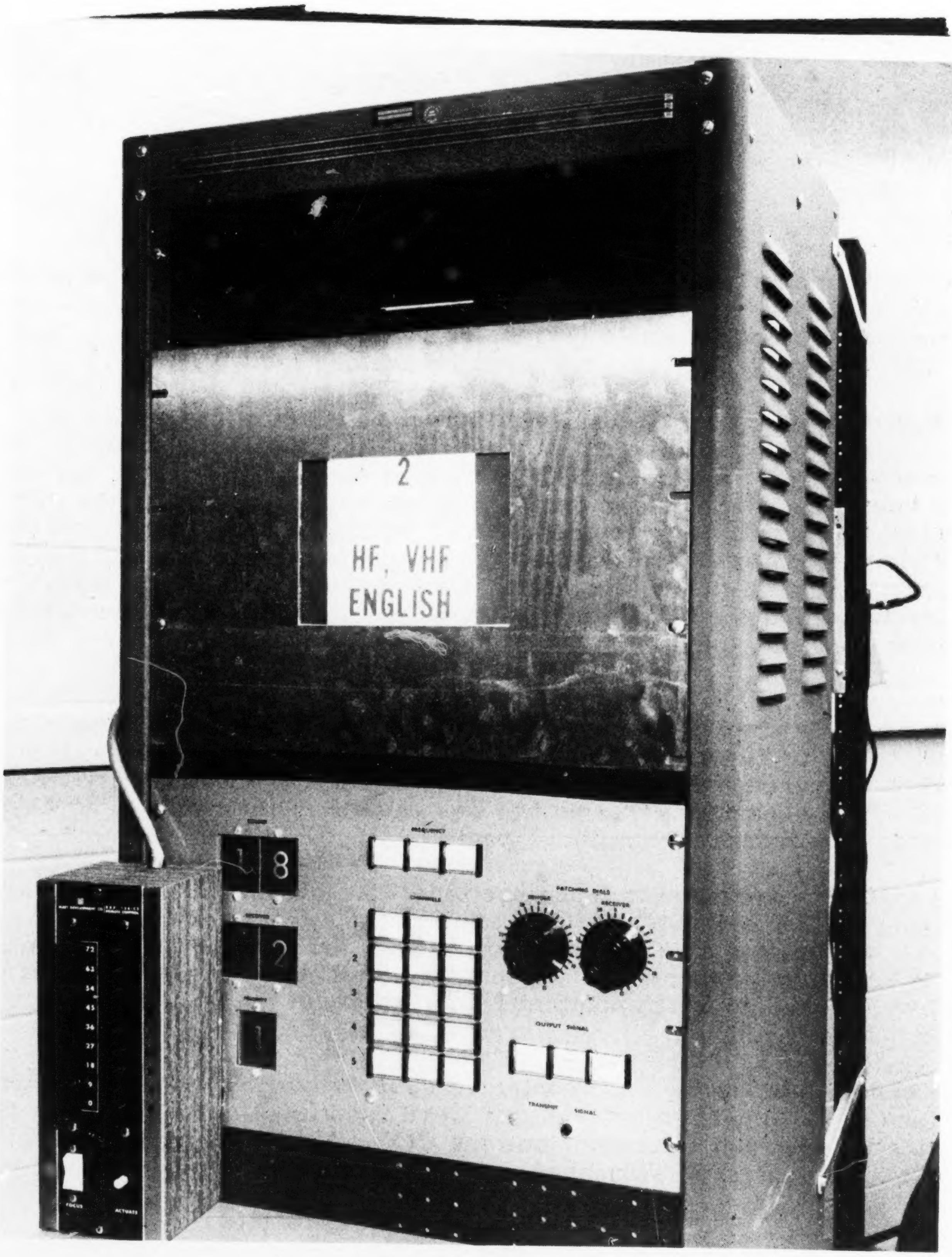

2
HF, VHF
ENGLISH
FREQUENCY
CHANNELS
1 8
2
1
1
2
3
4
5
72
63
54
45
36
27
18
9
0
OUTPUT SIGNAL
TRANSMIT
SIGNAL

Figure 1. Subject's Console

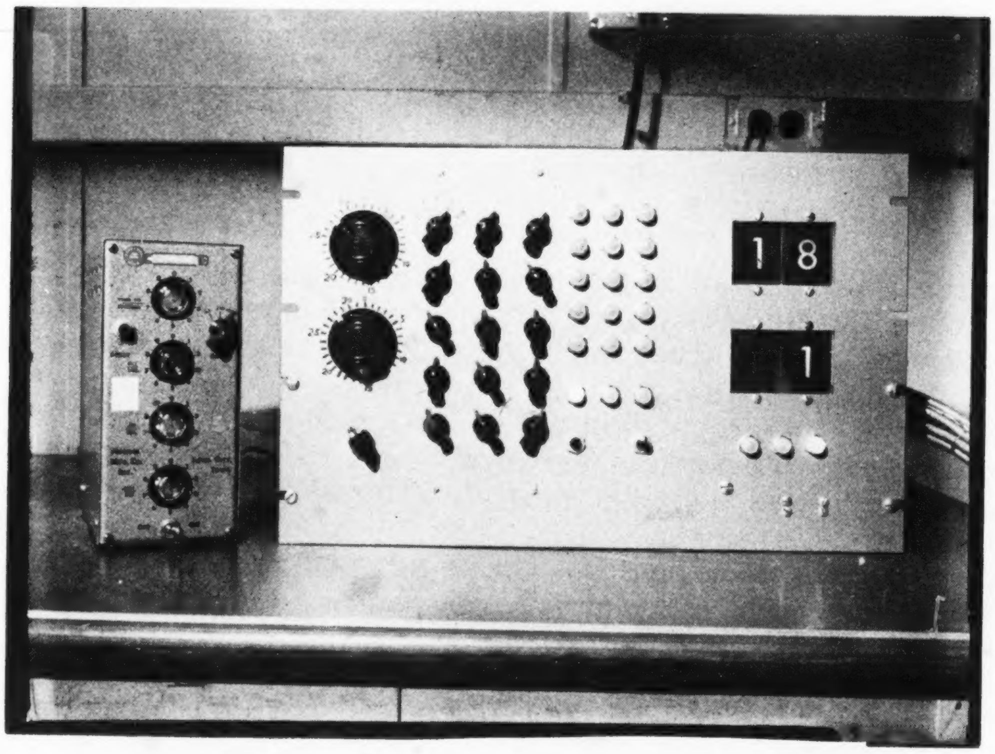


Figure 2. Experimenter's Console

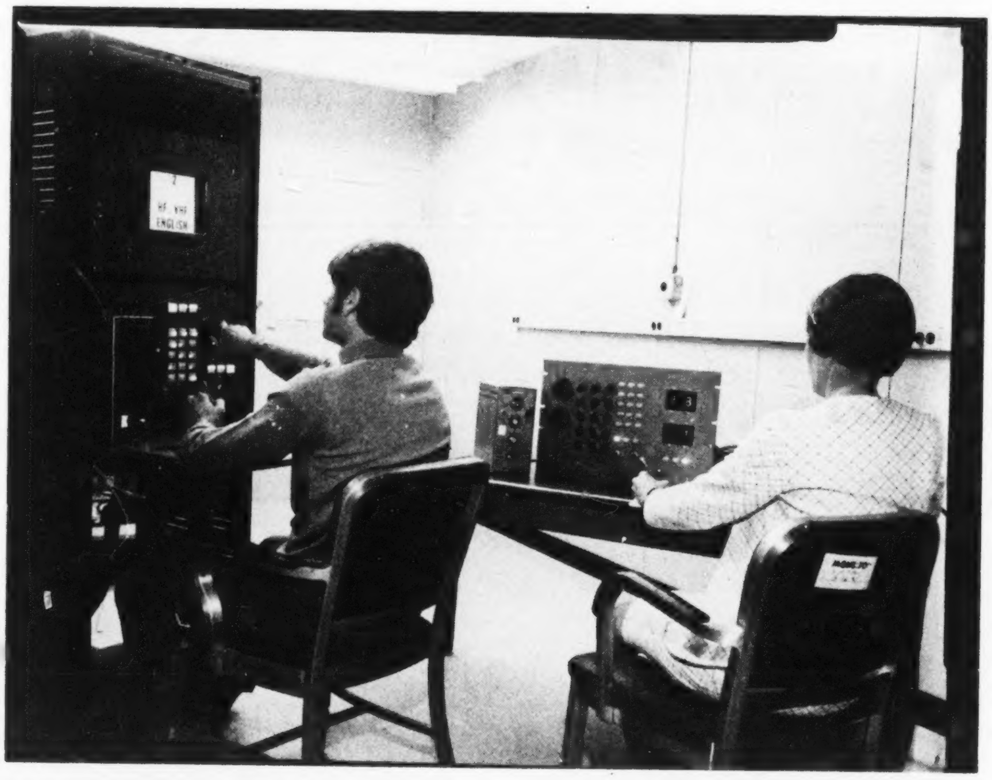


Figure 3. L, R, and T Console

- They indicated which channels were in use and the priority of the message of that channel. Message priority was indicated by the color of the illuminated indicator.*

- They were used to select a channel for transmission of messages to receivers.

The output section was located at the right side of the CCC. Two 32-position rotary switches, called "patching dials", were used to select the number of the sender and the receiver for each communication. Three microswitch switchlights were used for: (1) connecting the sender and receiver, (2) transmitting standby signals, and (3) transmitting busy signals. A momentary pushbutton, at the bottom right of the panel ("transmit signal" switch), stopped the timer, and thereby signalled the end of a subroutine.

Two sets of 30 problems were used. These sets were designed to be equivalent in terms of the principles and procedures used. They differed, however, in specific ways; i.e., in terms of the particular subscribers involved and the priorities of messages occupying communication channels. Within each set of 30 problems, 13 categories existed. These categories are listed in Table 2.

The problem sets were divided into five subsets, each containing six problems. Problems involved either one, two, or three subroutines. Each subset of problems totaled 12 subroutines. A Wollensak monaural tape recorder was used to provide instructions to the subjects. (See Bernstein and Gonzalez, 1969, for detailed instructions to subjects. These general instructions were the same for all groups and for all the experiments reported.)

Subjects

Ten male subjects ranging in age from 18 to 25 were used in the experiment. These volunteers were undergraduate students from the University of Minnesota and Hamline University. Subjects were paid on the basis of their performance on the task. Pay ranged from $2.50 to $4.00 per hour, in the same manner as the earlier L, R, and T experiments (Bernstein and Gonzalez, 1969).

Procedure

The subject's task was to arrange for the transmission of messages from one system subscriber (sender) to another (receiver). To provide better understanding of this task, a task analysis was performed (Bernstein and Gonzalez, 1969) for each of the operational sequences. The task analysis

---

*Note: Each switchlight was divided into four sections. Priority one messages were shown by the top two sections illuminated red, priority two was signalled by green in the lower left section, and priority three was a white indication in the lower right section.

Table 2. Types of Problems Used in L, R, and T Experiments I - V.

Problem Type	Number Presented
Subscriber to subscriber (S-S)	4
Subscriber to terminal to subscriber (S-T-S) due to language disparity	3
S-T-S -- due to frequency disparity	3
S-T-S -- due to frequency and language	2
S-T-S rather than interrupt	2
Interrupt (STBY)/SS	2
Interrupt (BUSY)/SS	2
Interrupt (STBY/S-T-S (language)	2
Interrupt (BUSY)/S-T-S (language)	2
Interrupt (STBY)/S-T-S (frequency)	2
Interrupt (BUSY)/S-T-S (frequency)	2
Deny request (STBY)	2
Deny request (BUSY)	2
Total	30

served as the basis for specifying the steps involved in ideal operation of the system and was helpful in organizing the instruction manual used during training. This manual was virtually identical to that used in Phase I.

Modes of Operation -- Although there were 13 categories of problems, only the following four basic modes of operation existed:

- Subscriber to subscriber (S-S)
- Subscriber to terminal to subscriber (S-T-S)
- Interrupt routine
- Deny sender's request to communicate

The selection of the appropriate routine was dependent on subscriber operating characteristics (language and operating frequency), priority of the requested transmission, and availability of channels. Each of the 25 subscribers could transmit or receive on one, two, or three frequencies; i.e., high frequency (HF), very high frequency (VHF), and ultra high frequency (UHF). Each subscriber was also associated with one of five languages: English, German, Greek, Spanish or Swedish. Language and frequency characteristics were provided on the 35 mm slides and could be selected for presentation on the information display.

All communications were conducted using either the S-S or S-T-S modes. In some cases, the interrupt mode was used but always in conjunction with either the S-S or S-T-S mode. Some problems involved a request under conditions which precluded the use of either the S-S or S-T-S modes. In these cases, the sender's request to communicate had to be denied.

S-S Mode -- When the sender and receiver used the same frequencies and the same language, and a channel was available on their common frequency, the S-S mode was used. To complete an S-S communication the following procedure was used:

(1) Select sender's number on the left "patching dial"

(2) Select receiver's number on the right "patching dial"

(3) Select the uppermost unoccupied channel on the least occupied common frequency

(4) Activate the "connect" switch

(5) Activate the "transmit signal" switch

S-T-S Mode -- The S-T-S routine was used under any of the following conditions: (a) Either a language disparity existed, (b) sender and receiver had different operating frequencies, or (c) there was an open channel within the receiver's capability, but not the sender's. The S-T-S routine was completed in two discrete parts. The first part included the following steps:

(1) Select sender's number on the left "patching dial"

(2) Select position 1 (terminal address) on the right "patching dial"

(3) Select a frequency within the sender's capability

(4) Activate the "connect" switch

(5) Activate the "transmit signal" switch

After completion of this subroutine, the message had been transmitted from the sender to the CCC. To relay from the CCC to the receiver, the following subroutine was employed:

(1) Select position 1 on the left "patching dial"

(2) Select the receiver's number on the right "patching dial"

(3) Select the uppermost open channel within the least-occupied frequency usable by the receiver

(4) Activate the "connect" switch

(5) Activate the "transmit signal" switch

Interrupt Routine -- The interrupt routine was used when all usable channels were occupied, and the requested transmission was of a higher priority than one or more of the messages occupying usable channels. In these cases the subject signaled the subscribers, using the to-be-interrupted channel that they must discontinue communication. If the interrupted channel contained a priority-two message, a standby signal was given, indicating that the interruption was temporary. An interrupted priority-three transmission was given a busy signal, indicating a longer-duration interrupt. The following steps constituted the interrupt routine:

(1) Select position 30 on both "patching dials"

(2) Select the channel to be interrupted

(3) Activate "standby" or "busy" for priority-two or priority-three transmissions, respectively

(4) Activate the "transmit" switch

(5) Wait for the channel to be cleared (indicated by the offset of the illuminated channel indicator)

As stated earlier, the interrupt routine was always used in conjunction with either the S-S or S-T-S routine. If the transmission was S-S, the channel was interrupted first and the S-S routine followed. However, when interruption was used with S-T-S, it occurred after the receipt of the message by the terminal and prior to the relay of the message to the receiver.

Deny Sender's Request to Communicate -- When all channels within the receiver's frequency capability were occupied with equal or higher priority messages than that currently requested for transmission, the request had to be denied. Requests were denied in the following way:

(1) Select position 1 on the left "patching dial"

(2) Select sender's number on the right "patching dial"

(3) Select any frequency usable by the sender

(4) Activate "standby" switch for priority one and two requests and the "busy" switch for priority three requests

(5) Activate the "transmit signal" switch

Test Sessions -- Subjects were given approximately one and one-half hours of training. The training session included the following phases:

(1) A tape recording of the manual was played to the subject.

(2) The experimenter presented five example problems to the subject and aided him, to the extent necessary, in the completion of the problems. The problems consisted of one each of the following types: S-S, S-T-S, S-S/Interrupt, S-T-S/Interrupt and deny sender's request.

(3) A 10-minute free study period was given.

(4) Thirty training problems were administered, with knowledge of results (KR) given after each problem. The KR was complete; i.e., any and all mistakes were pointed out and corrections were provided.

A trial (one problem) consisted of the following steps:

(1) The experimenter programmed the CCC.

(2) The experimenter said "ready".

(3) The CCC was illuminated by the experimenter, automatically starting the timer.

(4) The subject attempted to solve the problem. Each subroutine was timed separately. At the completion of one subroutine of a problem containing two or three steps, the timer stopped. The experimenter then noted the time and switched the console off, and then on again, restarting the timer. The subject then began the next step.

(5) KR was administered at the end of each problem.

Fourteen days after training, subjects returned for a transfer session. At the start of the second session, the subjects were informed that they would receive another set of 30 problems similar to those they had worked the previous day. No KR was given after these problems.

Half the subjects were given problem set A in training and problem set B at transfer, and the other subjects were given the two sets in reverse order. The order of subsets was counterbalanced across subjects. For example, subject 1 was given subsets in the order 1-2-3-4-5; subject 2 received order 2-3-4-5-1, etc. The order of the problems within subsets was randomized.

Experimental Design -- Table 3 shows the experimental design for this investigation. Five treatment groups were formed on the basis of the type of imagery used during training. Eight subjects were assigned to each group on a random basis. Sets A and B consisted of 30 problems each which were developed in Experiment I. Prior to training on the appropriate problem set, each group received 5 minutes of preliminary training as indicated in Table 3 in addition to 40 minutes of general instruction which was the same for all five groups.

Table 3. Experimental Design - L, R and T Experiment III

Group	Training Problem Set	Transfer Problem Set	Subject Number
I. Analogy Imagery Plus Text	A B	B A	1 - 4 5 - 8
II. Direct Imagery Plus Text	A B	B A	9 - 12 13 - 16
III. Direct Imagery	A B	B A	17 - 20 21 - 24
IV. Direct Text	A B	B A	25 - 28 29 - 32
V. Control - No Imagery No Text - Instruction Manual	A B	B A	33 - 36 37 - 40

Imagery Conditions -- The five imagery conditions used in this experiment are described as follows:

- Analogy imagery plus text -- This group received the cartoon slides which depicted the air-travel analogy from Experiment II (see Bernstein and Gonzalez, 1969). The slides were changed in that the balloons showing what the characters were saying, were taken out. This was done so that the effects of images and text would not be confounded. All text was contained in the script which was played to subjects from the tape recorder.
- Direct imagery plus text -- The direct imagery, like the analogy, was presented on 35 mm slides. However, rather than showing an analogous system, the direct conditions

showed cartoon representations of subscribers interacting with the CCC. Again, each of the four basic communication routines was represented. These slides are shown in Appendix B.

- Direct imagery -- These subjects were shown the direct imagery slides, but were not provided with the text describing each slide.
- Direct text -- Subjects in this group heard the text but were not shown the accompanying slides. Instead, subjects were told to form mental pictures of what they were hearing.
- Control -- Control group subjects were not given any imagery. The time normally used for imagery presentation was used by subjects to study the manual.

Scoring

Two measures of performance (i.e, response time and accuracy) were taken during the training and transfer sessions.

Response Time -- This measure referred to the elapsed time between the beginning and completion of a subroutine. If a subroutine was omitted, an elapsed time of 60 seconds was given.

Accuracy -- Accuracy data are expressed as percentages of subroutines correctly processed. The percentages were calculated according to a standard scoring method which deducted points from each subroutine for procedural errors. The more serious the error, the greater the scoring penalty. A maximum of five points was awarded for each subroutine. For example, an unnecessarily interrupted channel resulted in a 3-point penalty. . If the problem could have been accomplished by using a single S-S communication, involving only one subroutine, the subject's score for that problem would have been $\frac{(5-3)}{5}$ or 40 percent. If, however, the problem required two subroutines, the subject's score would have been $\frac{(10-3)}{10}$ or 70 percent. Likewise, a three-stop problem in which one three point error was made would have resulted in a score of $\frac{(15-3)}{15}$ or 80 percent.

RESULTS

The training and transfer performance curves for each of the five groups are shown in Figures 4 and 5. In general, the differences among treatment groups in time (Figure 5) and accuracy (Figure 4) were rather small.

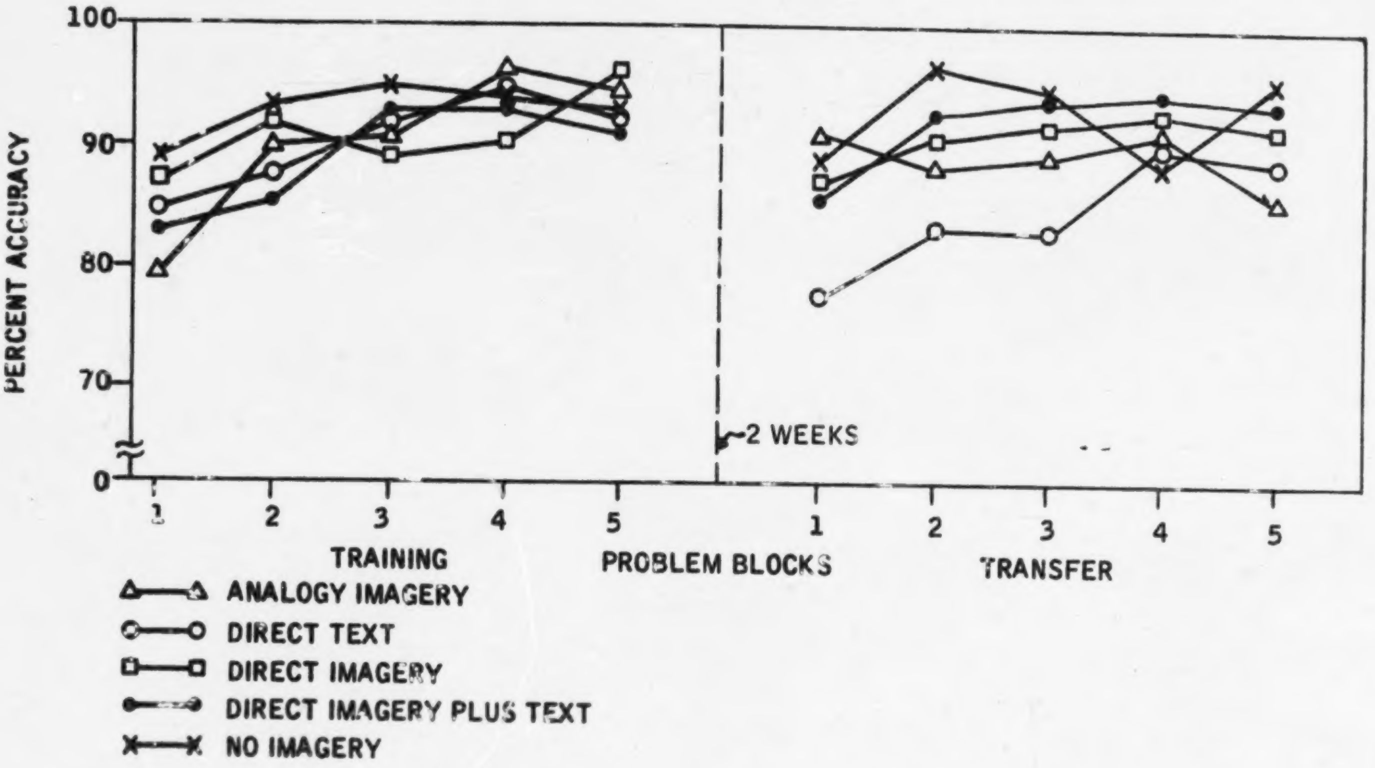


Figure 4. Percent Accuracy for Acquisition and Transfer Sessions

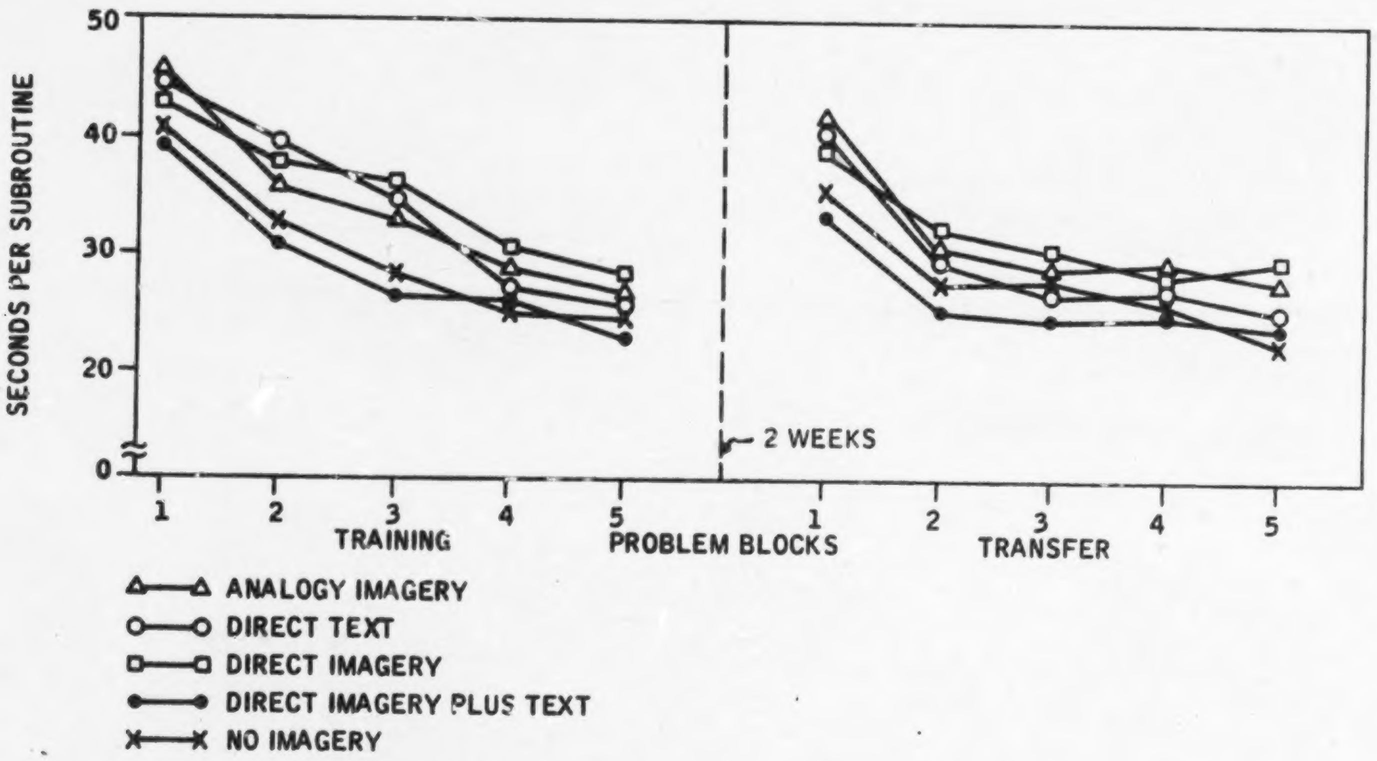


Figure 5. Mean Response Time for Acquisition and Transfer Session

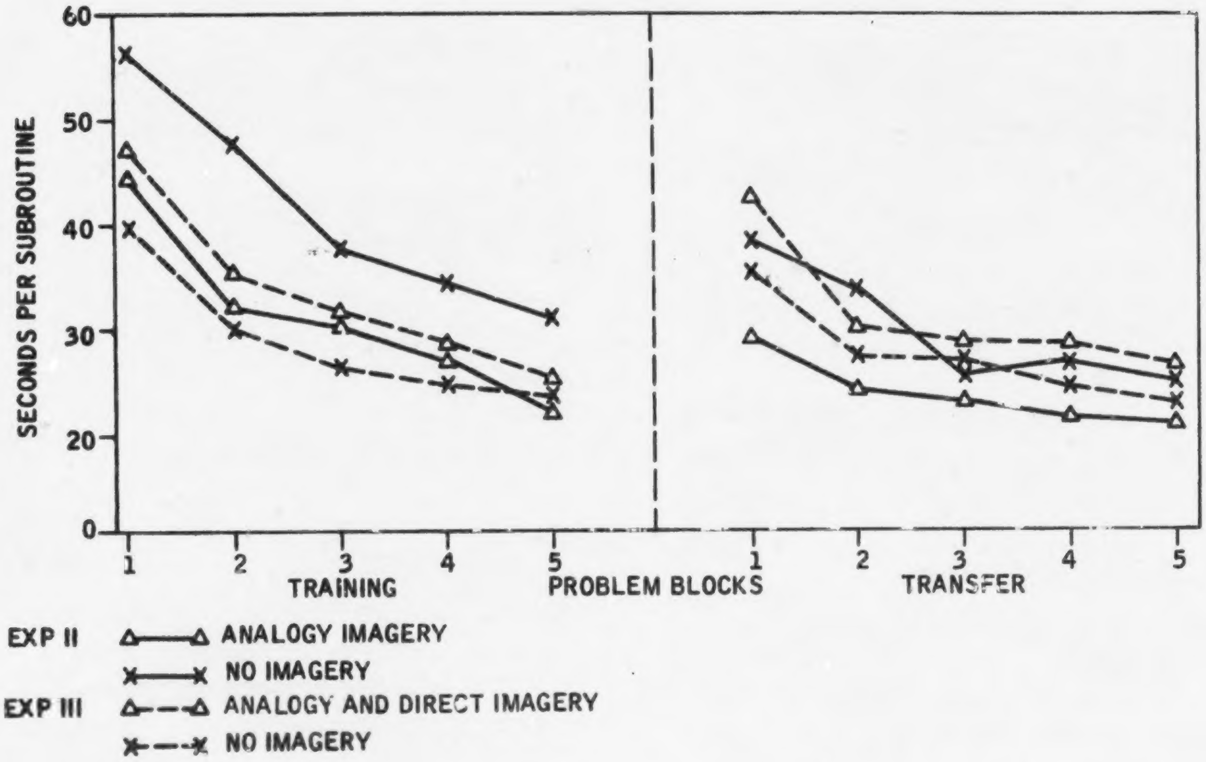


Figure 6. Mean Response Time for Acquisition and Transfer Sessions of Experiments II and III (Retention intervals for Experiments II and III were 7 and 14 days, respectively.)

Table 4. Analysis of Variance Summary - Response Times During Training Session

Source of Variation	df	MS	F
Imagery (U)	4	245.1	<1
Trial blocks (T)	4	1836.8	69.6**
(I)	35	377.4	---
I x T	16	18.0	< 1
T x S	140	26.4	---

$^{**}p < 0.01$

The overall response time performance of the imagery groups in this experiment are comparable to the analogy group of Experiment II (see Figure 6). However, there is a notable difference in the response times for the control groups of the two experiments. The No Imagery (Control) group of Experiment II performed substantially slower than the control group of Experiment III. A t-test was run on the total response times during training of these groups and was found to be statistically significant ($t = 4.2$, $p < 0.01$).

Analyses of variance (ANOVA's) (Tables 4 through 7) were performed on the acquisition and transfer data for both response time and accuracy scores for Experiment III. The various imagery conditions failed to produce a significant main effect in either session for either measure. The trial blocks effect was statistically reliable, showing a learning effect for both time and accuracy in both training and transfer sessions. A significant interaction was obtained between imagery and trial blocks for accuracy scores during transfer. This is shown by the non-parallel functions over trial blocks for the different experimental groups.

DISCUSSION AND CONCLUSIONS

The data of this study failed to support the imagery findings of Experiment II. The absence of an imagery effect seems to be associated with the higher performance of the control group (i.e., high fidelity - no imagery) of Experiment III relative to the control group of Experiment II. Had the no-imagery subjects of the previous investigation performed as well as the no-imagery subjects here, we would not have obtained a significant imagery effect in the earlier study.

The most obvious question posed by these results relates to the difference in performance of the control groups of two experiments. The likelihood that sampling error was responsible for the difference is very small ($p < 0.01$).

A more likely possibility is that differences in procedure accounted for the results. Such differences included the following:

1) In Experiment II, the instructions were read to the subjects by the experimenter. In Experiment III, the instructions were presented on audio tape.

2) Because more imagery slides were used in the direct conditions, an additional two minutes of free study was used in the Experiment III (seven minutes versus five minutes in Experiment II).

Table 5. Analysis of Variance Summary - Response Times During Transfer Session

Source of Variation	df	MS	F
Imagery (I)	4	217.6	1.0
Trial blocks (T)	4	1318.2	54.0**
S (I)	35	213.2	
I x T	16	12.7	<1
T x S (I)	140	24.4	

**p < 0.01

Table 6. Analysis of Variance Summary - Response Accuracy During Training Session

Source of Variation	df	MS	F
Imagery (I)	4	68.7	<1
Trial blocks (T)	4	502.8	12.1**
S (I)	35	167.2	
I x T	16	43.3	1.0
T x S (I)	140	41.6	

**p < 0.01

Table 7. Analysis of Variance Summary - Response Accuracy During Transfer Session

Source of Variation	df	MS	F
Imagery (I)	4	428	1.4
Trial blocks (T)	4	145.7	2.9*
S (I)	35	302.6	
I x T	16	88.1	1.8*
T x S (I)	140	49.5	

*p < 0.05

3) Training effectiveness may have increased because of the accumulated experience gained through conducting similar experiments by the experimenters in the present study. In particular, with more experience, the experimenter could better respond to subjects' questions.

4) Although irrelevant to the control group's performance, there was a change in the presentation of analogy images. The "balloons" used in Experiment II, containing messages by the figures in the cartoons, were removed for Experiment III. This was necessary to make the analogy cartoons equivalent to those used for the direct conditions. This change may have depressed performance for the analogy group in the present investigation.

Of these differences, it would seem that the most likely possibility is that the training procedure may have improved enough by the time this experiment was run to account for the reduced response time scores of the control group. Further, it may be that if training procedures are sufficiently effective, imagery adds little to performance.

SECTION IV

TECHNICAL APPROACH - EXPERIMENT IV

METHOD

Apparatus

The subject's task was similar to the previous studies. However, this investigation was conducted with the paper and pencil simulation of the equipment used in Experiment II (Bernstein and Gonzalez, 1969).

Subjects

Thirty male volunteer subjects from Introductory Psychology classes at Hamline University took part in the experiment. Subjects were paid on the same basis as in previous studies. Subjects were assigned to treatment conditions at random.

Procedure

Experimental subjects served one at a time. Before participating in the experiment, each subject was administered the Ammons Quick Test (QT) of verbal skills. In this test, the subject looks at a card containing four scenes. The experimenter reads a list of words, and as each word is presented, the subject identifies the scene which best illustrates the word. Each of the three cards has associated lists of 50 words, ranging progressively from "very easy" to "very difficult". Subjects are scored on the basis of total words correct. For each subject, the test took 10 - 15 minutes to administer and about a minute to score.

Subjects scoring above the mean for college students (QT scores of 125 were average) were placed in the high verbal skills group, and those scoring below 125 were placed in the low verbal skill group. Of the eight subjects who scored exactly 125, four were arbitrarily assigned to the high and four to the low verbal skills group.

Subjects in each group were assigned randomly to one of the three imagery groups: analogy, direct, or control (no imagery).

The experimental session began immediately after each subject was assigned to an imagery group.

Conduct of the experimental session was essentially the same as in the previous low-fidelity (paper and pencil) experiments, except that there was no transfer session. Half of the subjects in each group were tested with problem set A and half were tested with problem set B, as in the previous experiments.

RESULTS

The performance curves for Experiment IV are shown in Figures 7 through 10.

Figure 7 shows the response time data for high and low verbal skills groups. Figure 8 shows performance response time measures for the three levels of imagery, i.e., analogy, direct, and control (no imagery).

Figures 9 and 10 show the accuracy score data in percent correct for high and low verbal skills and for the three imagery conditions.

Analyses of variance (ANOVA's) were performed for both response time and accuracy data. Summaries of these analyses are presented in Tables 8 and 9. Neither the verbal skill levels nor the imagery conditions produced a statistically significant main effect. The trial blocks effect was again statistically significant ($P < 0.01$) for both the response time and accuracy data.

DISCUSSION AND CONCLUSIONS

This investigation was conducted for two reasons: (1) To help resolve the discrepancy in results between Experiments II and III, and (2) to test the hypothesis that the effectiveness of imagery interacts with verbal skills.

The findings of the current study are consistent with that of Experiment III, showing no evidence that imagery improves performance as the operation of the synthetic communication system. This was true regardless of the verbal skill level of subjects. as measured by the Ammons QT in this experiment. It should be pointed out that the imagery conditions in Experiments III and IV and that of Experiment II were different in that the verbal "balloons" were removed in Experiments III and IV.

Once again, the control group of the present experiment outperformed its counterpart of Experiment II. The hypothesis that overall training effectiveness had improved after Experiment II seems to be further supported by these data. It appears that subjects were able to comprehend the functional relationships existing among elements in the system without the presentation of either the direct or analogous imagery.

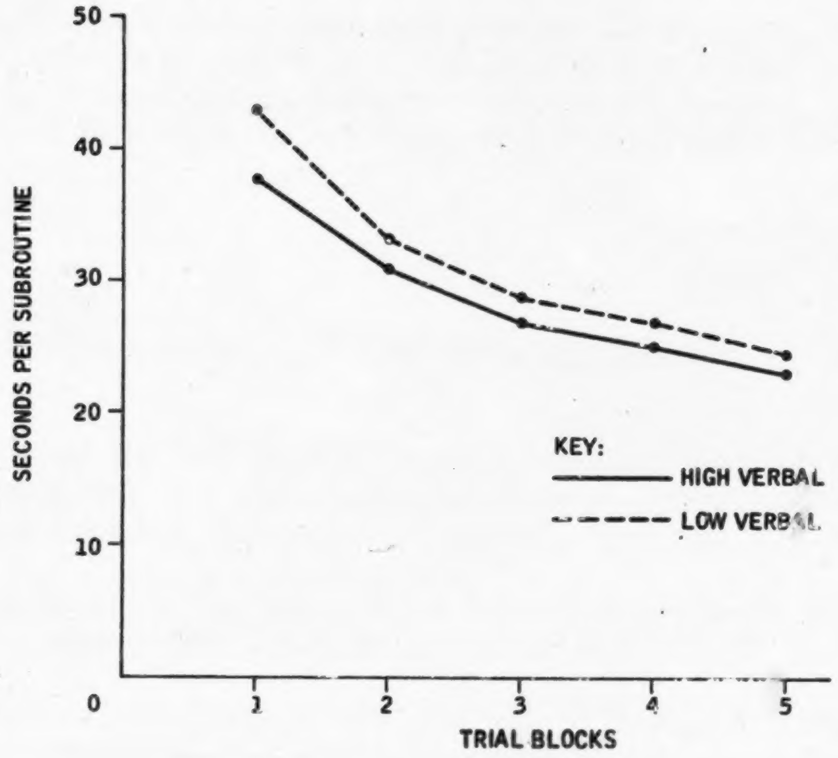


Figure 7. Mean Time to Complete Subroutines (Verbal Skill Groups)

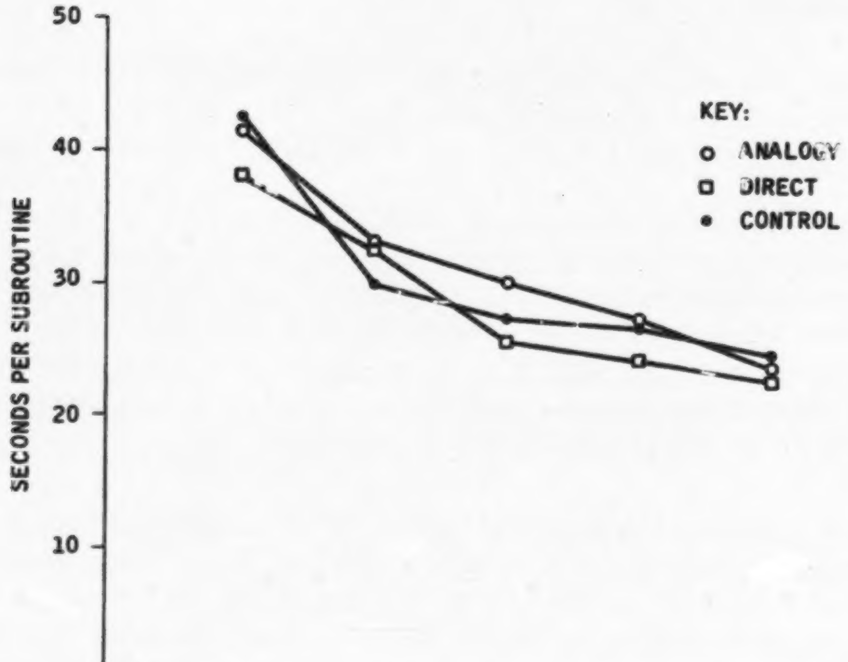


Figure 8. Mean Time to Complete Subroutines (Imagery Levels)

44

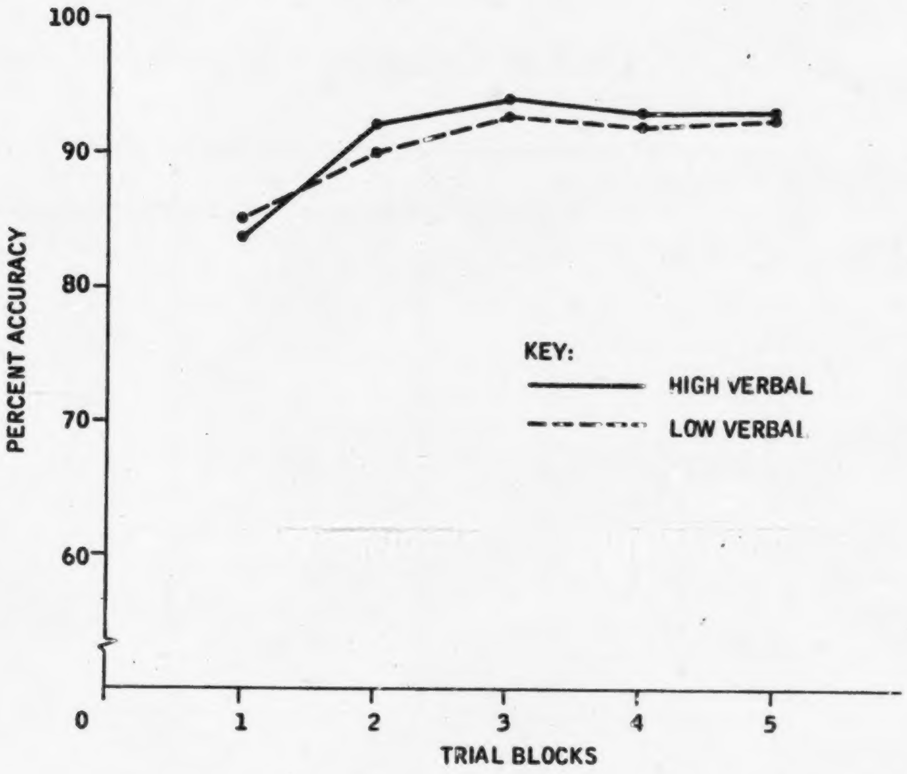


Figure 9. Mean Percent Accuracy

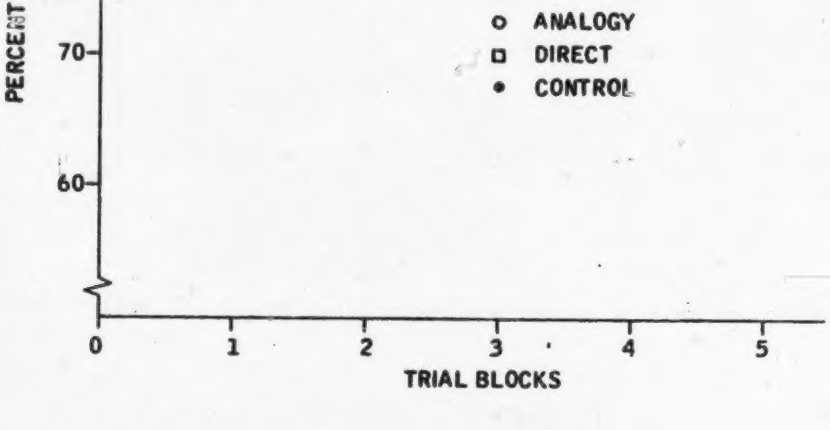


Figure 10. Percent Accuracy

Table 8. Analysis of Variance Summary - Response Time During Training

Source of Variation	df	MS	F
Verbal skill (V)	1	399.8	1.25
Imagery (I)	2	136.5	<1
Trial Blocks (T)	4	1264.8	33.78**
V x I	2	29.7	<1
V x T	4	28.5	<1
I x T	8	17.3	<1
S (VI)	24	319.6	
V x I x T	8	36.7	<1
T x S (VI)	96	37.4	

** P<0.01

Table 9. Analysis of Variance Summary - Response Accuracy During Training

Source of Variation	df	MS	F
Verbal skill (V)	1	14.6	<1
Imagery (I)	2	135.2	1.50
Trial Blocks (T)	4	242.1	6.99**
V x I	2	37.1	<1
V x T	4	17.0	<1
I x T	8	57.8	1.67
S (VI)	24	90.0	
V x I x T	8	25.0	<1
T x S (VI)	96	34.6	

** P<0.01

The hypothesis that imagery should be more effective for subjects with lower verbal skills was not supported by these data. There was no evidence that high versus low scorers on the Ammons QT were different in their ability to learn to operate the communication system. Stated differently, the Ammons test failed to predict performance on the criterion task.

The absence of predictive validity for this test may be the result of an excessively restricted range in verbal skills for the subjects used. This statement is supported by the fact that over half of the subjects (16) scored at the mean for college students or within two points of the mean.

The present investigation does not appear to provide conclusive results regarding the effects of imagery on subjects of varying verbal skills. It may be worthwhile to run another group of subjects who score significantly lower than the college students used in this study. The subjects of this experiment could then be considered high scorers and the new group, low scorers. The additional group should then be subdivided according to imagery condition in a manner similar to that used with the subjects already run.

If an interaction could be demonstrated between imagery and verbal skill, it would help resolve the discrepancy in findings between Experiment II and Experiment III and IV. That is, the difference in performance associated with imagery in Experiment II, may have been the result of less than adequate instruction. Because the training was so dependent on verbal instruction, the imagery may have been effective when the verbal presentation was sub-optimal. A second approach to this problem may be used. An experiment in which quality of instruction is treated factorially with imagery may also explain the inconsistency of the obtained findings.

However, these two suggested approaches are closely related. Instructions may be effective for people of relatively high verbal skill and ineffective for lower skilled subjects. Further research certainly seems indicated relating to verbal skills, quality of instruction, and imagery.

SECTION V

TECHNICAL APPROACH - EXPERIMENT V

METHOD

Apparatus

Training equipment for this study was of the paper and pencil type. Three versions of the training equipment were used (Figures 11 through 13). These versions varied systematically in terms of their similarity to the operational hardware. The highest fidelity representation (Figure 11) was identical to that used in Experiments II and IV. The medium fidelity drawing (Figure 12) differed from the operational equipment in that the front panel components were reoriented by a 90-degree shift. The lowest fidelity drawing (Figure 13) contained a 90-degree shift plus a change in all controls and displays.

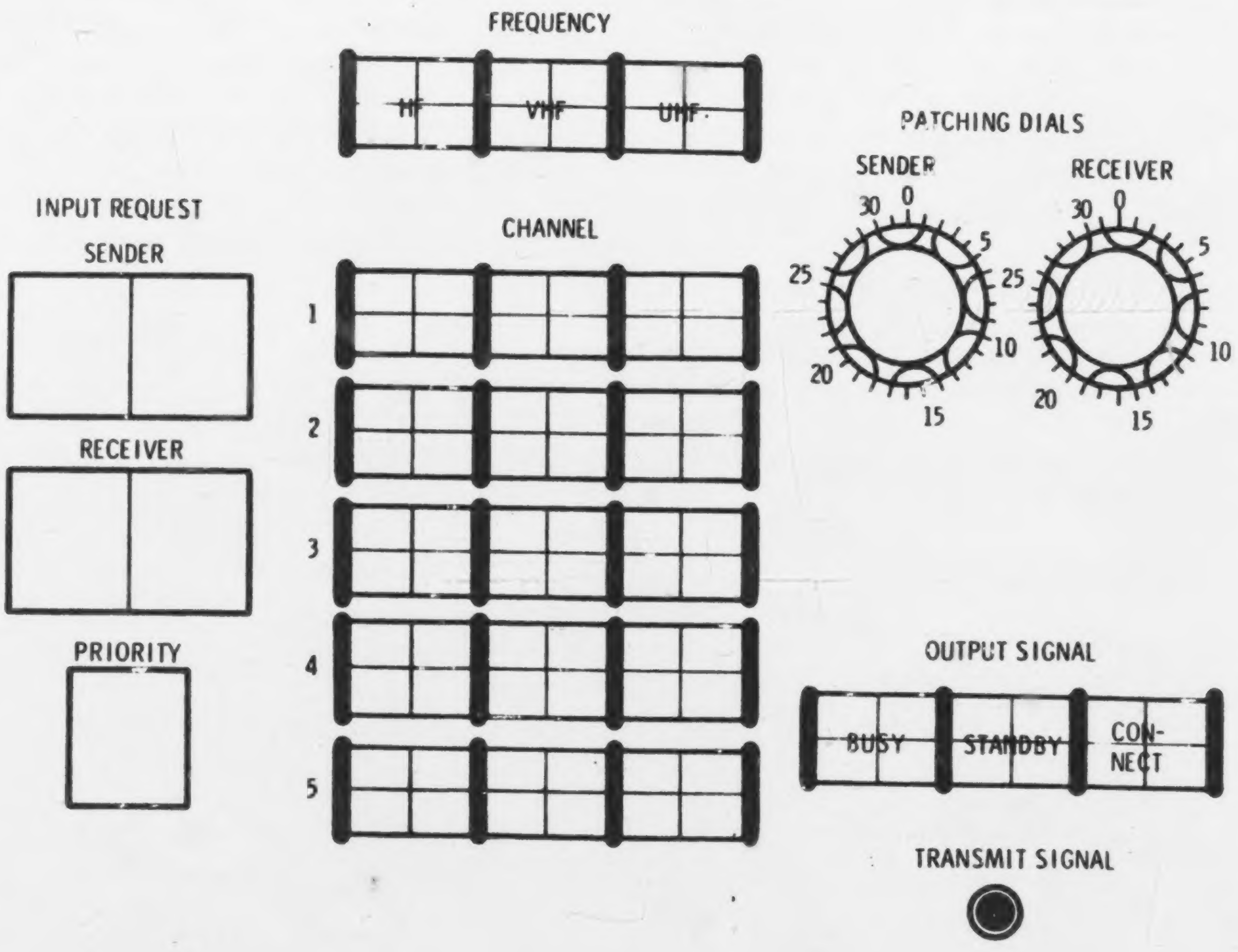

FREQUENCY
HF
VHF
UHF
PATCHING DIALS
SENDER
RECEIVER
30
0
5
25
10
20
15
INPUT REQUEST
SENDER
CHANNEL
1
2
3
4
5
RECEIVER
PRIORITY
OUTPUT SIGNAL
BUSY
STANDBY
CON-NECT
TRANSMIT SIGNAL

Figure 11. High-Fidelity Panel

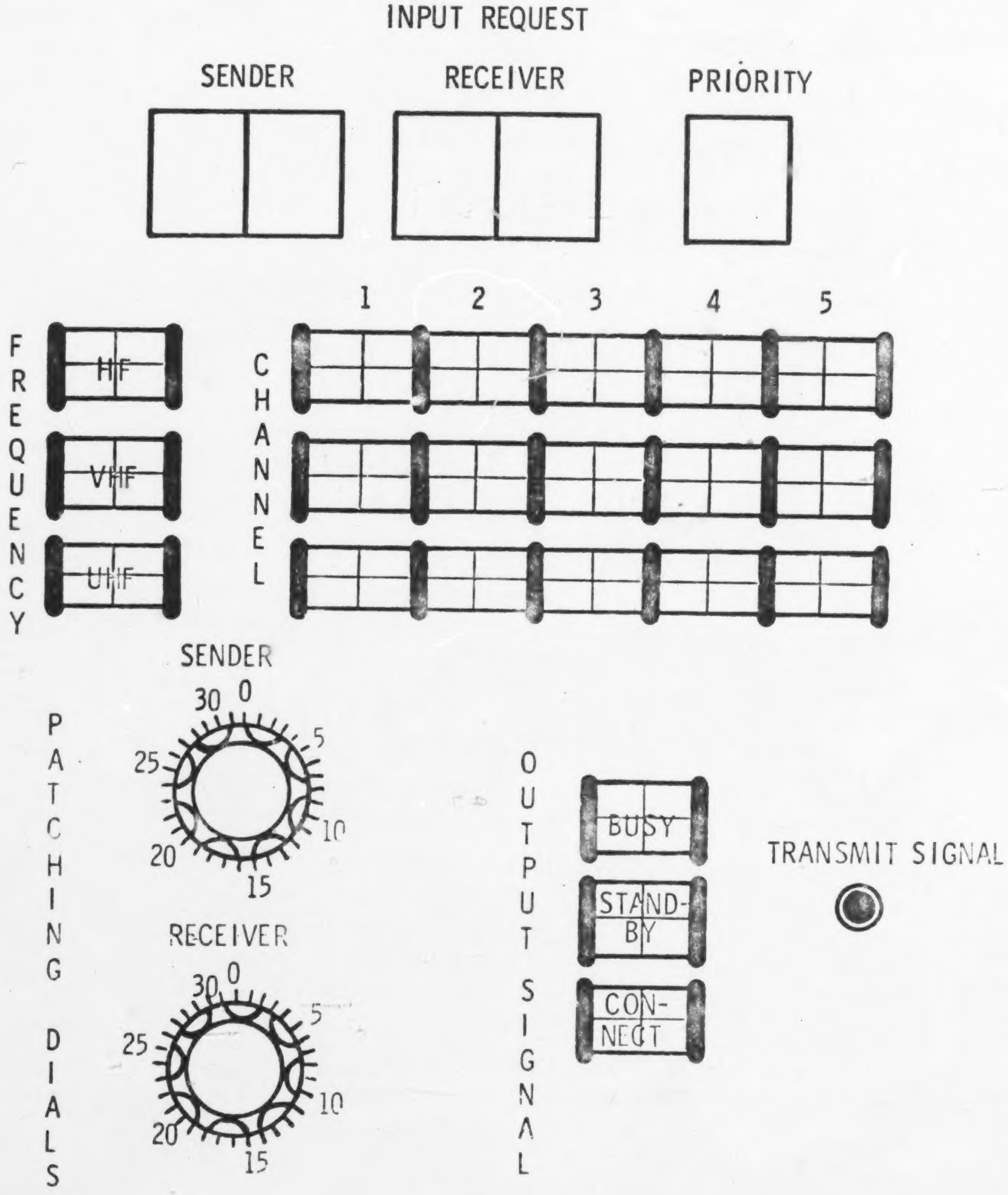
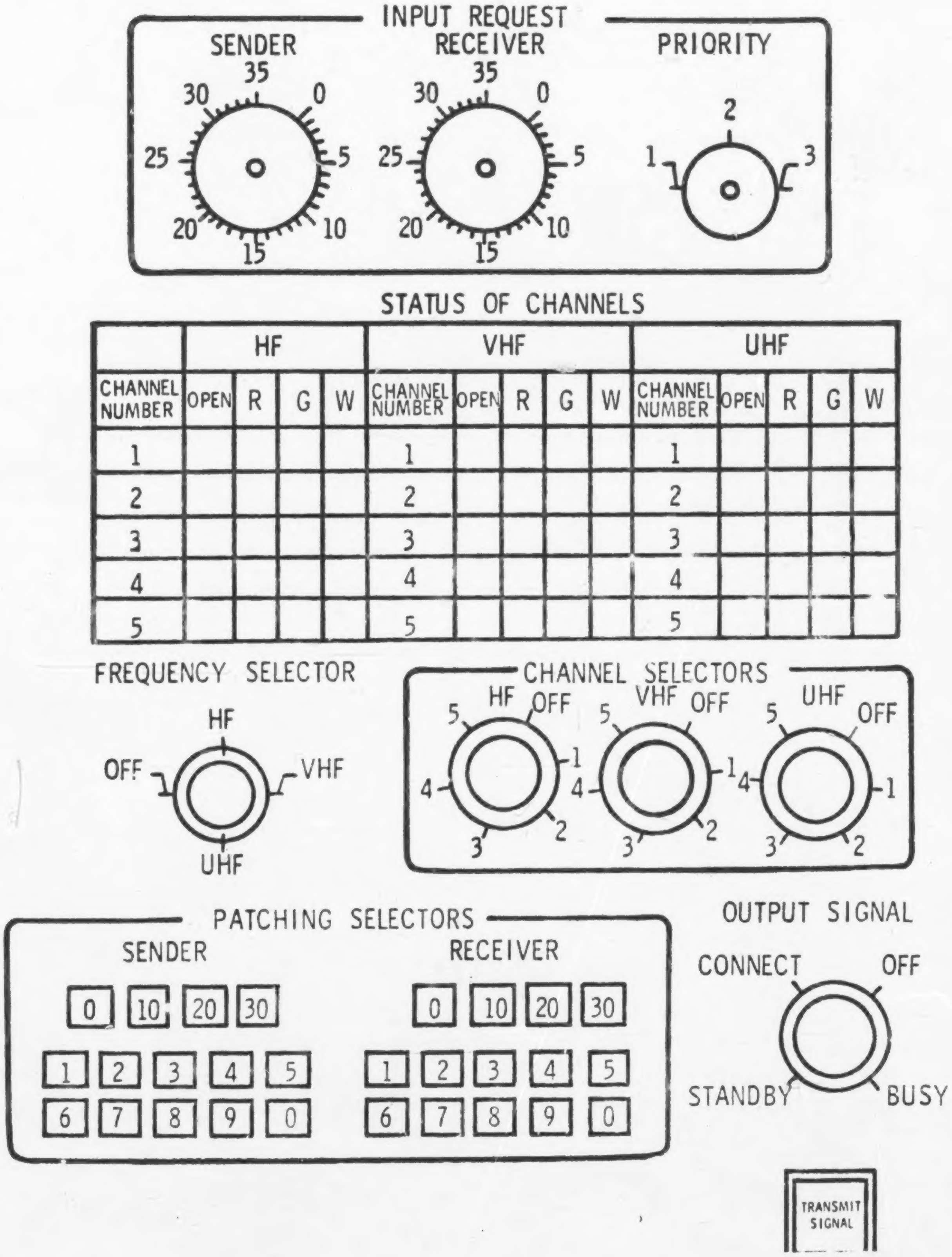

INPUT REQUEST
SENDER
RECEIVER
PRIORITY
FREQUENCY
HF
VHF
UHF
CHANNEL
1
2
3
4
5
PATCHING DIALS
SENDER
RECEIVER
0
5
10
15
20
25
30
OUTPUT SIGNAL
BUSY
STAND-BY
CON-NECT
TRANSMIT SIGNAL

Figure 12. Medium-Fidelity Panel


INPUT REQUEST
SENDER
35
30
0
25
5
20
10
15
RECEIVER
35
30
0
25
5
20
10
15
PRIORITY
2
1
3
FREQUENCY SELECTOR
HF
OFF
VHF
UHF
CHANNEL SELECTORS
HF
OFF
5
1
4
3
2
VHF
OFF
5
1
4
3
2
UHF
OFF
5
1
4
3
2
PATCHING SELECTORS
SENDER
0
10
20
30
1
2
3
4
5
6
7
8
9
0
RECEIVER
0
10
20
30
1
2
3
4
5
6
7
8
9
0
OUTPUT SIGNAL
CONNECT
OFF
STANDBY
BUSY
TRANSMIT SIGNAL

STATUS OF CHANNELS

	HF				VHF					UHF				
CHANNEL NUMBER	OPEN	R	G	W	CHANNEL NUMBER	OPEN	R	G	W	CHANNEL NUMBER	OPEN	R	G	W
1					1					1				
2					2					2				
3					3					3				
4					4					4				
5					5					5				

Figure 13. Low-Fidelity Panel

Subjects

Forty-eight male students from the University of Minnesota and Hamline University took part in the experiment. Subjects were assigned to the six treatment conditions at random. The subjects were again paid from $2.50 to $4.00 per hour for their participation.

Procedure

The procedure for this study was the same as that used previously, except that all subjects were trained with a paper and pencil representation of the equipment.

Again, one set of problems was used for training and the other for transfer. Differences in fidelity occurred during training, with all subjects transferring to the actual equipment. The transfer session was conducted 14 days after training.

Experimental Design

Table 10 shows the design for this study. Six groups of subjects were used in a 2 x 3 factorial arrangement. Two levels of response fidelity high

Table 10. Experimental Design - L, R and T - Experiment V

Training		Transfer	Subject Number
Stimulus Fidelity	Response Fidelity		
Low	Verbal	Problem Set A	1 - 4
		Problem Set B	5 - 8
	Written	Problem Set A	9 - 12
		Problem Set B	13 - 16
Medium	Verbal	Problem Set A	17 - 20
		Problem Set B	21 - 24
	Written	Problem Set A	25 - 28
		Problem Set B	29 - 32
High	Verbal	Problem Set A	33 - 36
		Problem Set B	37 - 40
	Written	Problem Set A	41 - 44
		Problem Set B	15 - 48

(written) and low (verbal), and three levels of stimulus fidelity (high, medium, and low similarity) were used[1]. Eight subjects were assigned to each treatment group.

RESULTS

The response time data (Figures 14 and 15) suggest that variations in stimulus and response fidelity had little effect on training effectiveness. Response times during transfer show an effect similar to that obtained in Experiment II; i.e., a marked increase in speed from the first to the second block of problems.

The accuracy data (Figures 16 and 17) also show little effect of levels of stimulus and response fidelity. Performance curves during transfer, for written versus verbal groups, are very nearly identical. Subjects who were trained under the lowest stimulus fidelity condition were somewhat less accurate at transfer than were medium and high-stimulus fidelity groups (Figure 16).

Apparently, accuracy performance suffered a greater loss at transfer than did response time. Although subjects quickly regained their speed, accuracy consistently remained below first session levels.

Analyses of variance were performed on the time and accuracy scores from acquisition and transfer sessions. These ANOVA'S are summarized in Tables 11 through 14. Neither method of responding nor stimulus fidelity produced significant main effects. Only one interaction was found to be significant ($p < 0.05$), viz., stimulus fidelity x trials in percent accuracy during training; no interpretation for this finding is available at present.

DISCUSSION AND CONCLUSIONS

The results of this investigation extend the findings of Experiment II. Considered together with the other research in the area, we now have considerable evidence that relatively high levels of training effectiveness for procedure tasks can be maintained in the absence of high physical fidelity between training and transfer tasks. This appears to be true both in terms of the stimulus and response components of the task.

These findings lend support to Wittrock's (op. cit.) concept of mediated generalization as the basis of transfer in procedural skills. Apparently, subjects can cope with variations in the stimulus and response components of the task as long as the basic system structure and functions remain unchanged.

---

[1] The written response is considered to be of higher fidelity than the verbal response because the former requires hand-equipment relationships which are similar to those involved in operating the actual equipment whereas such hand-equipment relationships are absent with the verbal response. These definitions, however, are arbitrary and unique to this situation.

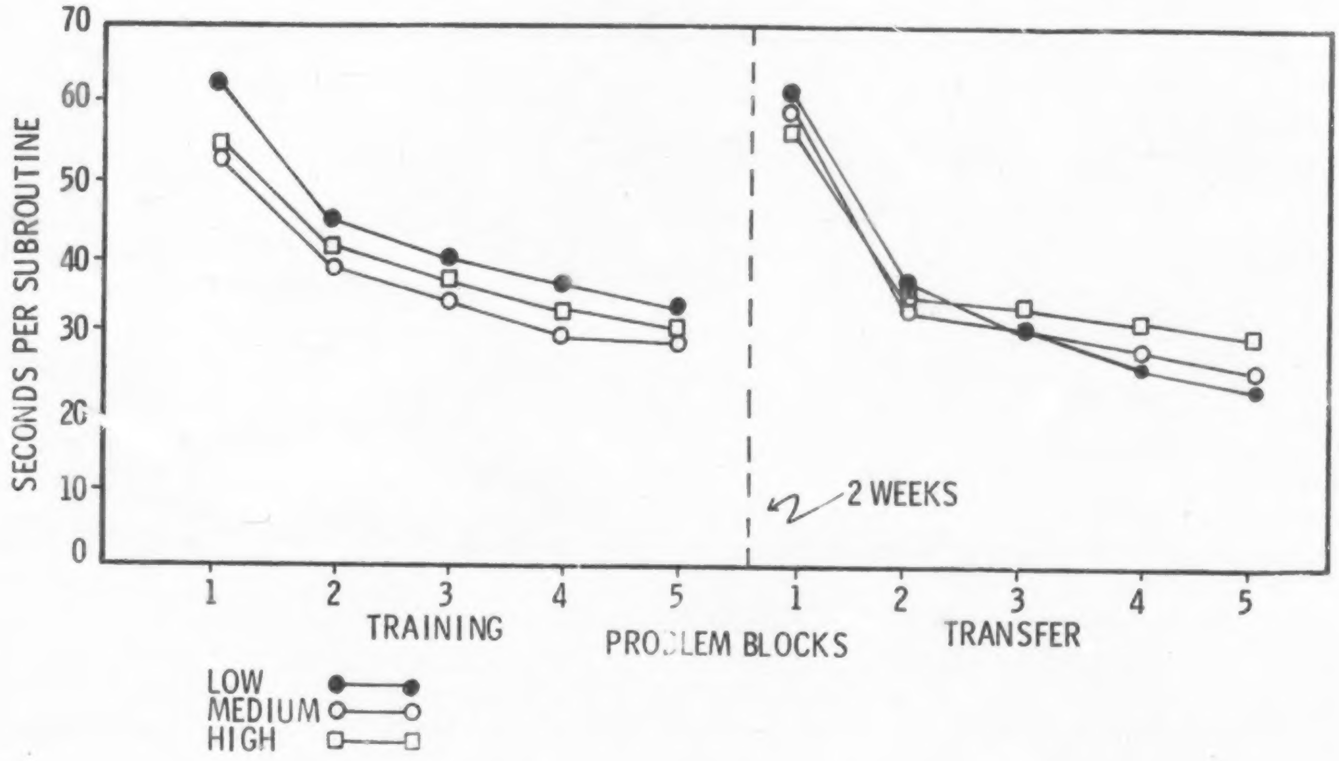


Figure 14. Mean Response Time as a Function of Stimulus Fidelity

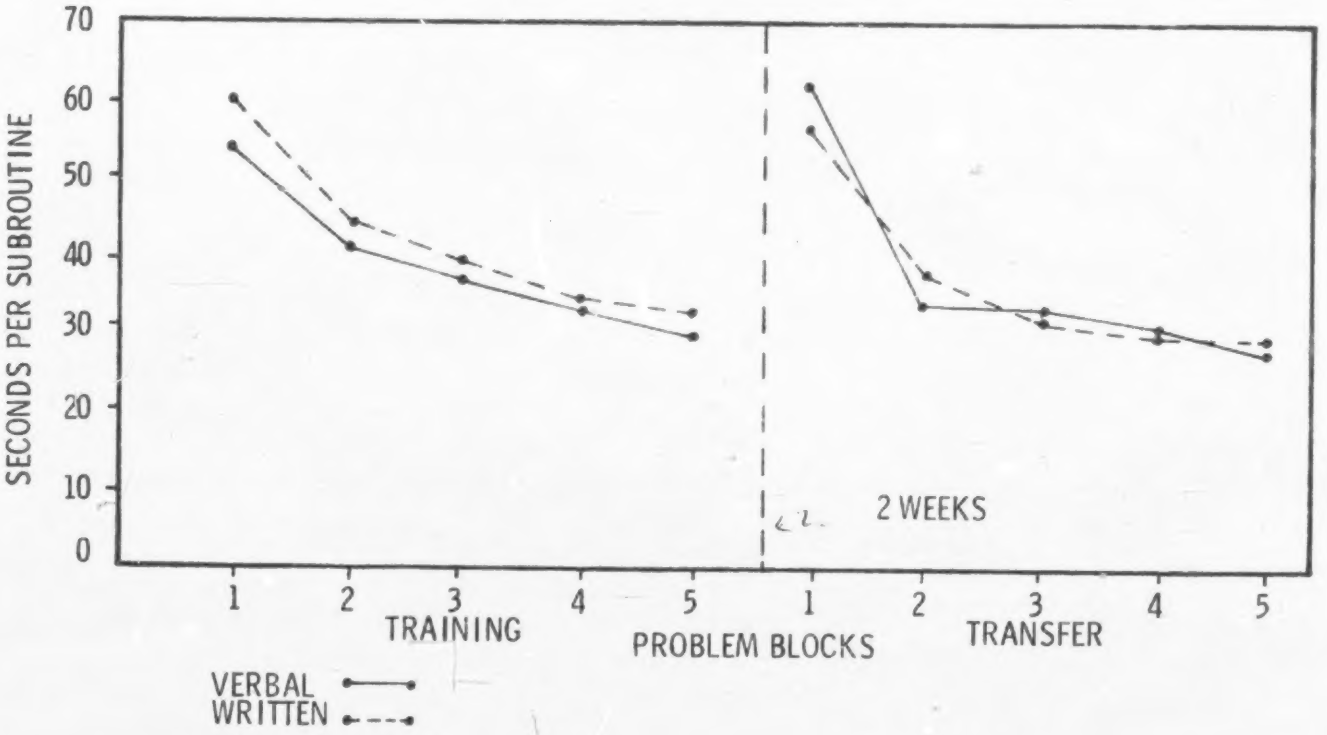


Figure 15. Mean Response Time as a Function of Response Fidelity

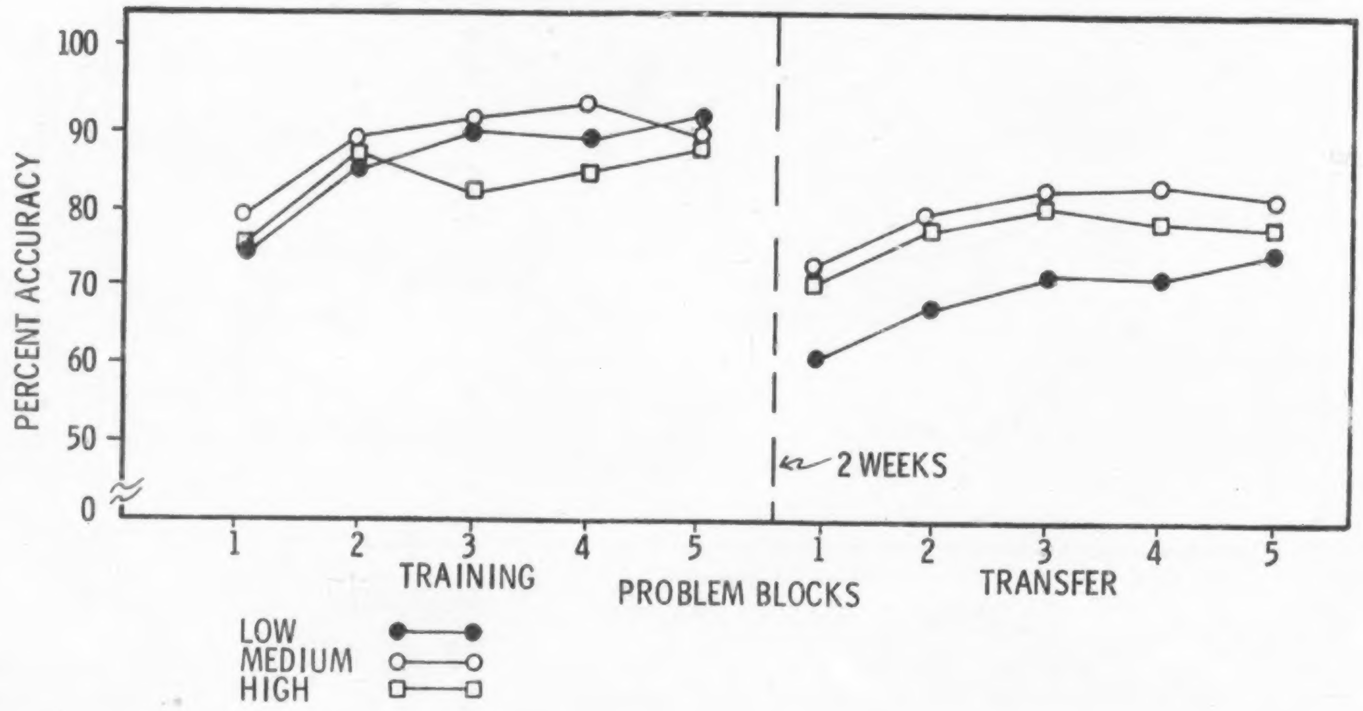


Figure 16. Accuracy as a Function of Stimulus Fidelity

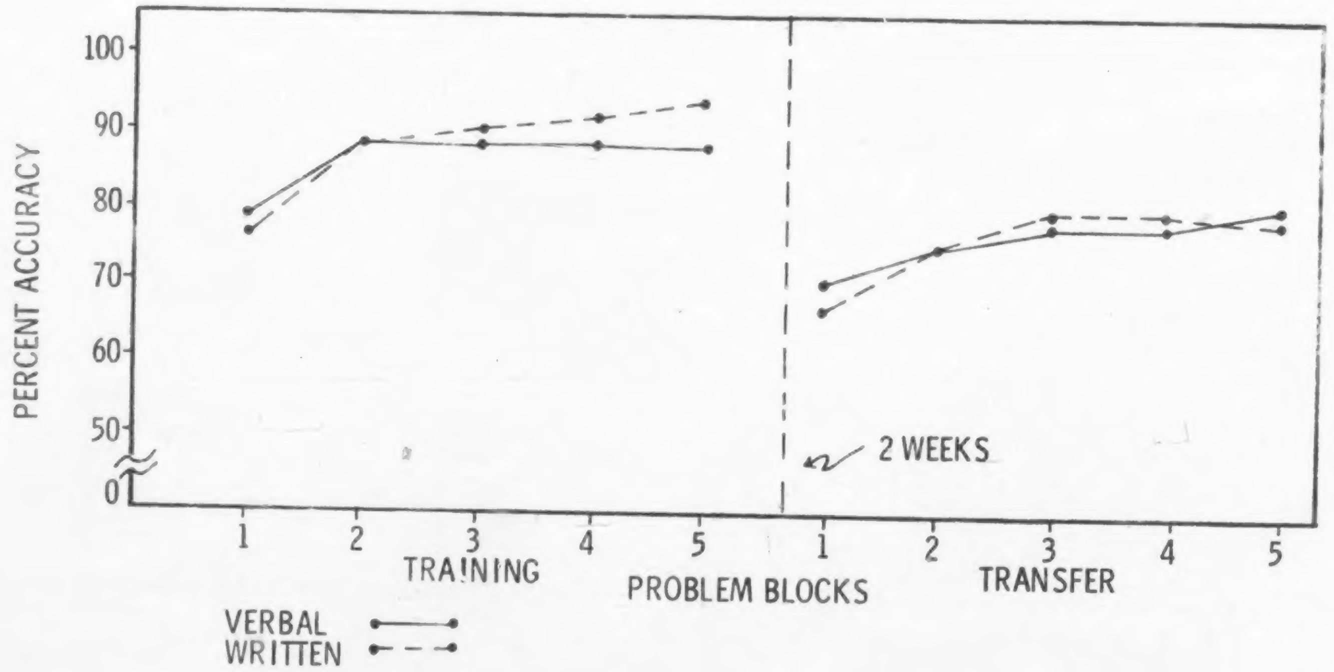


Figure 17. Accuracy as a Function of Response Fidelity

54

Table 11. Analysis of Variance Summary - Response Time During Training

Source of Variation	df	MS	F
Response method (R)	1	729.4	< 1
Stimulus Fidelity (F)	2	847.5	< 1
Trials (T)	4	4958.8	99.99**
R x F	2	301.5	< 1
R x T	4	59.8	1.37
F x T	8	31.1	< 1
S (R x F)	42	921.9	
R (R x F x T)	8	42.0	< 1
T x S (R x F)	168	49.6	

** $p < 0.01$

Table 12. Analysis of Variance Summary - Arcsin Transformed Percent Accuracy Scores During Training

Source of Variation	df	MS	F
Response method (R)	1	0.1337	< 1
Stimulus Fidelity (F)	2	0.2235	1.516
Trials (T)	4	0.6271	21.56**
R x F	2	0.3182	2.06
R x T	4	0.0526	1.80
F x T	8	0.0607	2.08*
S (R x F)	42	0.1541	
R (F x T)	8	0.0399	1.37
T x S (R x F)	168	0.0291	

* $p < 0.05$
** $p < 0.01$

Table 13. Analysis of Variance Summary - Response Time During Transfer

Source of Variation	df	MS	F
Response method (R)	1	2.6	< 1
Stimulus Fidelity (F)	2	56.7	< 1
Trials (T)	4	7983.4	54.37**
R x F	2	24.7	< 1
R x T	4	171.5	1.54
F x T	8	84.5	< 1
S (R x F)	42	993.6	
R x F x T	8	102.1	< 1
T x S (R x F)	168	146.9	

** $p < 0.01$

Table 14. Analysis of Variance Summary - Arcsin Transformed Percent Accuracy Scores During Transfer

Source of Variation	df	MS	F
Response method (R)	1	0.0003	< 1
Stimulus Fidelity (F)	2	0.4345	< 1
Trials (T)	4	0.2482	10.85**
R x F	2	0.1109	< 1
R x T	4	0.0091	< 1
F x T	8	0.0216	< 1
S (R x F)	42	0.4849	
R x F x T	8	0.0167	< 1
T x S (R x F)	168	0.0229	

** $p < 0.01$

It is worth noting, however, that subjects' performance on early trials of the transfer task reflects changes to the equipment and different methods of responding. Subjects require more time to process communication routines on these initial trials. However, response times improved very rapidly by the second block of trials.[2] As was mentioned in the discussion of the Phase I results, unless initial performance is critical, high physical fidelity does not appear to be a stringent requirement.

The discrepancy between time and accuracy measures, in terms of transfer performance, was an interesting finding. The accuracy data obtained in this experiment tended to be poorer than that obtained earlier. This was not the case for response time. Accuracy at transfer was not reduced in Experiment III to the extent that it was in the present investigation, suggesting that the two-week retention interval was not solely responsible. On the other hand, accuracy scores did not differ significantly as a function of fidelity conditions. Moreover, this discrepancy was not apparent in Experiment II. Thus, it does not seem likely that low-fidelity training is the primary reason for the lower accuracy scores. Possibly, a combination of a two-week retention interval and training with paper and pencil simulations produced the obtained effect.

Also, it should be kept in mind that knowledge of results was not provided during transfer trials. Therefore, an initial drop in accuracy would tend to perseverate throughout the transfer session. An experiment appears to be needed in which feedback is treated as an independent variable during transfer, with training provided under high- and low-fidelity conditions.

Integrative Review of Imagery and Fidelity of Simulation

An effort is made in this section to consider the fidelity and imagery characteristics of the present experiments in relation to other research in the area. Although the literature review contained in this section might typically be presented in the introduction, this section contains frequent references to aspects of the experiments which suggests that it be more appropriately presented here.

---

[2] That this effect is influenced by the transfer to the actual equipment from paper and pencil training is suggested by a similar result obtained in a Phase I experiment (Experiment II). That study showed a decrement in performance at the start of transfer and rapid subsequent improvement for a group trained with paper and pencil, while subjects trained on the actual equipment showed considerably less initial loss. Thus, since the initial transfer effect was due to paper and pencil training in this previous experiment, it seems reasonable to assume that the similar effects in the present experiment were also a result of pencil and paper training (although a control group trained on the actual equipment is not available for comparison in the present case). An implication of this assumption is that the difference between training on the actual equipment and training with paper and pencil has a greater effect on initial transfer than differences in either the configuration of the controls and displays or variations in the response mode employed. (This point is indicated by the close clustering of all curves at all stages of transfer in Figures 14 and 15.)

Imagery -- The imagery variable manipulated in the present studies represented a calculated gamble; namely, the attempt to extend recent thinking on the role of imagery (I) in verbal learning to the specialized use of pictures in an instructional situation. Laboratory research on imagery has concentrated on development of objective quantitative measure of I (e.g., Paivio, Yuille, and Madigan, 1968) and on the functional status of I in studies of adults' and children's verbal learning. Scholarly reviews of this research are not wanting, and include summary articles by Paivio (1969, 1970), Reese (1970), Rohwer (1970), and Bower (in press). The context of such efforts has been primarily that of "list learning" in its various forms (i.e., PA learning, serial list learning, free recall, recognition, and memory span techniques). Such studies obviously emphasize the S's retrieval of specific information, scored in terms of discrete words on word groupings, rather than the S's ability to perform conceptual and procedural tasks. Though the emphasis on "discrete" words in such laboratory tasks should not be confused with a lack of concern for conceptual and relational principles, it is true that the task of the present experiments departs considerably from the typical list learning laboratory test. Likewise, the psychological scaling of single words for properties of I, concreteness-abstractness, and meaningfulness finds no ready application to the instructional methods evaluated in the present research.

These distinctions between laboratory "list learning" and laboratory-instructional investigations regarding the role of imagery in verbal learning need not pose an insurmountable obstacle to future efforts to take advantage of knowledge concerning imagery processes in instructional or training programs. Rather, the available literature suggests a number of clues (albeit by hindsight) for the subsequent design of such research, and these leads are discussed briefly below.

According to Paivio (e.g., Paivio and Yarmey, 1966) verbal stimuli may arouse sensory images which will mediate response recall (i.e., the conceptual-peg hypothesis). Accordingly, Paivio (1969) has shown that meaningfulness and imagery have separable effects, with that of I on the processing of the stimulus term in PA learning. Meaningfulness has its greatest effect on the response term. Future work on procedural-instructional programs might thus concentrate efforts on developing vivid word instructions and pictorial representations of stimulus or cue-function material, with less emphasis on imagery for the procedural response component.

Pictures versus Words -- The attempt to facilitate performance through pictorial representations of key concepts during the instruction session may be traced to its laboratory based antecedents fairly directly. According to Paivio (1969), "Objects or their pictorial representations arouse non-verbal images directly and, to the extent that such images facilitate the formation of associative connections with response members, pictures should be even more effective than concrete nouns, at least as stimulus members" (p. 254). Studies conducted prior to 1966 had examined the effect of pictures versus words on the stimulus or response side in PA learning, though Paivio and Yarmey (1966) were the first to investigate words versus pictures in all possible combinations (S-R) in a 2 x 2 factorial design. Consistent with the conceptual-peg hypothesis, Paivio and Yarmey found PA learning with pictorial stimuli as cue terms to be superior to verbal stimulus conditions.

Most of the picture-versus-word research has been performed with PA tasks, though some imagery research on this topic has also been done with free recall (FR) and serial learning procedures. And, in FR, a number of studies (Lieberman and Culpepper, 1965, Paivio, Rogers, and Smythe, 1968) have shown higher levels of recall for familiar objects than for the word names of the objects.

Paivio (1970) has recently reviewed a series of studies comparing pictorial and sentence (verbal context) depicting of pair relations. This work is a step closer to the present efforts. In his review, Paivio suggests that both imagery and verbal symbolic processes are operative in S's learning of pair relations (citing only Milgram, 1967, with a negative finding on imagery). Davidson and Adams (undated) studied children's learning of associations between pictorial representations of objects, and found learning aided by presentation of the objects in relation to one another (e. g., a rope around the neck of a jar) compared to simple juxtaposition of the objects (e. g., rope and jar side-by-side). In a related study directly concerned with imaginal processing, Bower (in press) investigated associative learning of paired words representing objects. One group was instructed "to imagine the two objects" (represented by words) "one at a time separated in their imaginal space, like two pictures being seen on the opposite walls of a room". The other group was instructed to imagine scenes in which the two objects interacted (in a manner similar to that explicitly used by Davidson and Adams). In cued recall, the "interactive" group recalled 71 percent of the responses that were to have been associated with the cues, while the "separate" imagery condition recalled far fewer items (46 percent).

Such experiments provide obvious leads for future work on aiding instructional routines with either pictorial representations or deliberate imaginal procedures. For example, the Ss in the present study were not explicitly instructed to process the pictorial aids in any particular way, or to relate the events within any single cartoon one to another. Yet, recent laboratory efforts of the sort described above suggest that development of associative relations between cue and response words or objects is very much a function of the interactive processing which may occur, rather than just a simple consequence of using pictures themselves. In this context, facilitation of retention results not itself from imagery, but from "integrated" imagery of the sort mentioned above. That is, the integration is achieved by the setting of to-be-learned materials in certain specific contexts (e. g., pictures of related objects, sentences relating key words, etc.), or directing S to process the materials himself in like fashion. Thus, Rohwer (1966) and Paivio (1969) find associative learning of pairs of common nouns to be facilitated by an intervening verb or preposition, but not aided by an intervening conjunction. The latter presumably does no more to establish an integrative imagery or verbal context effect than merely listing the terms separately.

Is this last factor one possible (though unintended) characteristic of the use of pictures to benefit verbal instruction in the present experiments, and hence a contributing element to the generally null imagery results here obtained? Although an exact answer is impossible as yet, the possibility of an affirmative answer is strong enough for the integrative hypothesis to receive more attention in future investigations.

Several investigations have been performed to compare potentially differential cue properties of words and pictures outside the "list learning" context, and these seem deserving of mention relative to the work reported above. Basically, such studies offer leads on the degree to which words and pictures may arouse sufficiently different responses to suggest caution in the instructional combining of particular verbal and pictorial representations. Deno, Johnson, and Jenkins (1968) investigated the similarity of associations evoked by pictures of objects and the verbal labels designating those objects. The purpose of the study was to gain some psychological purchase on the finding that word and pictoral stimuli may produce differential effects in tasks of learning and retention (e. g. , Deno, 1968; Jenkins, Neale, and Deno, 1967; Kale, Grosslight, and McIntyre, 1955). The authors found that in terms of associative meaning (Deese, 1962), "a picture does not represent the same concept as a word" even though the two are nominally in correspondence (i. e. , the word "apple" and a picture of an apple, etc.). Similar findings have been obtained by Bourisseau, Davis, and Yamamoto (1965) and Anderson (1966) in terms of such logical dimensions as "sense" impressions and bipolarity ratings in which responses to word and picture stimuli were scored. The same problem has been investigated from a quite different perspective by Shepard and Chipman (1970), who were attempting to tie down the concept of "internal representation" for pictorial and verbal stimulus events. The S's task was to give judgments of rated similarity for the shapes of fifteen states, based upon (a) only the state names, or (b) the visual state outline without the name. While Conrad (1964) found stimulus similarity (and hence confusion errors in short-term memory) to be based upon the acoustic rather than the visual properties of the stimuli, "suggesting that the internal representation of visually presented letters was at least initially related to the names of the letters," Shepard and Chipman found just the reverse. Namely, Ss appear to rate similarities among shapes cued by names much as they do when presented with the actual visual outlines. This type of inquiry seems fairly basic to an understanding of some of the processes underlying imagery for verbal cues, and the authors comment that such efforts are really in their infancy.

Effect of Pictures on Comprehension -- The literature regarding the effects of pictures on text comprehensions is in principle fairly closely related to the experimental manipulations performed during the current project. However, such work does not offer the attractive theoretical base as does that of imagery (though hindsight again suggests the role which imagery may play in some of the comprehension studies). A serious problem for valid generalization is that much of the instructional research on pictorial aids to comprehension has been performed with children rather than adults. With these cautions in mind, two reports in this area must be noted. Koehnke (1968) studied Ss' ability to identify the main idea of a paragraph of prose material with content-relevant pictures as the independent variable. The Ss were third and sixth graders and four treatments were evaluated, as follows: (1) text only; (2) text plus pictures, but no instructions to use the accompanying pictures; (3) text plus pictures, minimal directions on pictures use; (4) text plus pictures, "maximum" directions to use the pictures. Koehnke found that pictures did not add to Ss' ability to state the main ideas of the paragraphs, even with maximum directions as to the use of the pictorial material.

Structured Presentation Modes -- Bower (1970) provides an excellent review of research on grouping and relating of to-be-learned verbal materials. Three points clearly applicable to the present studies are discussed below. First, Bower describes a wide range of variables which may influence the extent to which Ss more effectively organize and learn verbal materials, of which I is but one of many potentially influential factors in recall. Second, Bower reaffirms the importance of distinguishing between E and S codes in verbal learning, with special emphasis upon E's discovering and making effective use of the latter in the design of the material presentation. (Note also in this context such diverse applications of S codes as those described by Tulving, 1968, Boltwood and Blick, 1970; Montague, Adams, and Kiess, 1966; and Montague and Wearing, 1967). In the present context, the pictorial stimuli used must be classified, at least for the time being, as E codes. Hence, we do not know their relevant stimulus properties as processed by S. Instead, the pictures represented the investigators rendition of an effective I stimulus related to the critical instructional material. Were the studies to be redone, it might well be desirable to generate the pictorial aids on the basis of S's descriptions of the images aroused by the actual spoken and written instructional text.

Finally, Bower describes several experiments concerned with the role of natural conceptual hierarchies in learning and retention of large sets of words. In one experiment, 28 to-be-learned words were ordered into a natural conceptual hierarchy constructed from a structural rule. In the control condition, the 28 words were presented in a random manner. Differences in recall favoring the structural-rule organization were striking; more words were recalled on the first trial for the "organized" condition than on the last trial by the Ss receiving randomly structured materials. This effect may be mediated by visualization processes, but it seems premature and possibly inaccurate to invoke imagery itself as a determinant factor. Instead, the experiments demonstrate the considerable value of careful and systematic organization of verbal materials according to conceptual rules. In the present context, the requisite verbal recall may have been facilitated by its own structural organization in the absence of any overt pictorial representation. This hypothesis is subject to test in research designed to deliberately manipulate both the textual and pictorial properties of the instructional material and should have high priority in future LRT instructional research.

Fidelity of Simulation -- Gagne (1962) provides an excellent overview of simulators in terms of instruction and on-line training objectives. Campbell (in press), in an Annual Review chapter devoted to Personnel Training and

Development, finds relatively few studies explicitly concerned with fidelity of simulation (where he expected to find many). A good many service supported studies of flight simulation have been carried out (e.g., Ellis, Lowes, Matheny, and Norman, 1968; Demaree, Norman, and Matheny, 1965) but appear outside the specific criterion-task concerns of the present research.

In a study of radar training requirements, Silver, Jones, and Landis (1965) give some empirical support to the idea that paper and pencil simulation devices (as used in the present studies) may successfully be employed for highly skilled operators (though not for task-naive ones). Research on individual differences in perceptual-motor skills (e.g., Fleishman, 1967) has consistently demonstrated that the ability requirements of a task may change over periods of practice. The question then arises: what components of the task, and at what level of task proficiency, should be simulated? The present research was concerned primarily with task simulation for the relatively early stages of practice. Clear answers regarding effects of degrading fidelity of simulation at initial levels of task proficiency are somewhat inconsistent, and made more so by potential retention X fidelity interactions, as reported by Grimsley (1969) for guided missile trainees, and by Hammerton and Tickner (1967) for Ss learning remote-control operations. Briggs and Naylor (1965) varied degree of task organization and fidelity of the training task (a real radar scope versus a cardboard cutout), and found team training not to transfer well when the training simulation was of low fidelity. Briggs and Johnson (1966) subsequently found stimulus fidelity more important than response fidelity in this situation (a finding with some interesting ties to the earlier description of I effects given above).

While the studies described in this report show that E has considerable latitude in permissible variations of the display representation, training on a paper mock-up is not identical to similar training on the CCC (note Figures 16 and 17). However, as Gagne (1962) points out, fidelity or degree of simulation may meaningfully refer to the specification of tasks to be performed by S during training, rather than to the formal resemblance of training and criterion equipment.

Simulation may not only involve the systematic training of certain task components while other situational variables are deliberately excluded; but may also involve, for example, variations of operational stimulus materials. The criterion task used in the present studies (the CCC) allowed the simulation of training mode features, but did not involve stimulus materials (messages) identical to those encountered in typical Naval tasks. Most simply, there are many shades of meaning ascribed in the use of the term simulation, only one of which has yet been systematically explored in this report.

SECTION VI

TECHNICAL APPROACH - RETENTION EFFECTS

A major problem with research investigating learning, retention and transfer is the size required of experiments in which all three processes are studied. Any single L, R & T experiment is automatically a three-factor investigation. Various classes of variables are associated with learning, retention, and transfer. For example, imagery was treated as a variable associated with learning, and fidelity of simulation has been considered a transfer variable.

The most obvious variable, relating to retention, is the time elapsing between training and transfer sessions. We have not manipulated this variable within any one experiment for two basic reasons. First, retention interval, at least in the present context, can be studied only as a between-group variable. Thus, the size of an experiment is multiplied by the number of retention intervals studied. Second, retention was not considered a high priority variable by participants at the L, R & T technical meetings -- probably because forgetting curves have been obtained frequently in the past and look similar regardless of the conditions under which they have been obtained. Unless one is interested in the retention characteristics of a particular task, it is usually sufficient to know that the forgetting curve will be a monotonic, decreasing function of time.

Retention data are available in the present program by examining the retention interval across experiments. Because a control group (high fidelity - no imagery) was run in each of the first three experiments, assessment of retention interval effects is possible.

Figures 18 and 19 show these data, expressed in percent savings scores. Experiment I provided the one-day values, Experiment II the seven-day interval, and Experiment III gave the 14-day scores. Each group contains two curves. One is for savings between block 1 of the training session and block 1 of the transfer session. The other curve shows savings averaged across blocks of trials. One graph (Figure 18) shows savings scores based upon time data, and the other (Figure 19) shows savings data based upon accuracy scores.

$$\frac{\text{Time}_{\text{Acq.}} - \text{Time}_{\text{Transfer}}}{\text{Time}_{\text{Acq.}}} \times 100$$

For the data in Figure 19, savings was calculated as

$$\text{Savings} = \frac{\text{Percent Errors}_{\text{Acq.}} - \text{Percent Errors}_{\text{Transfer}}}{\text{Percent Errors}_{\text{Acq.}}} \times 100$$

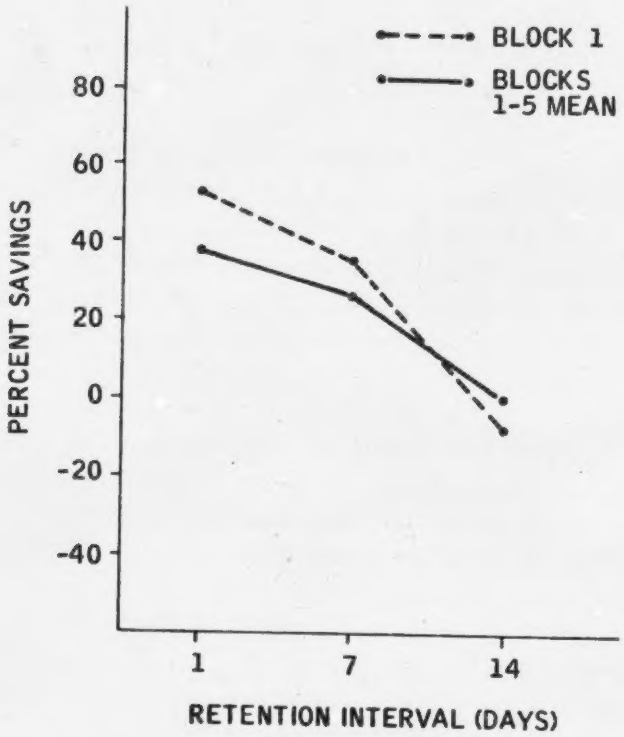


Figure 18. First Block and Average Savings Based on Response Time Scores, as a Function of Retention Interval

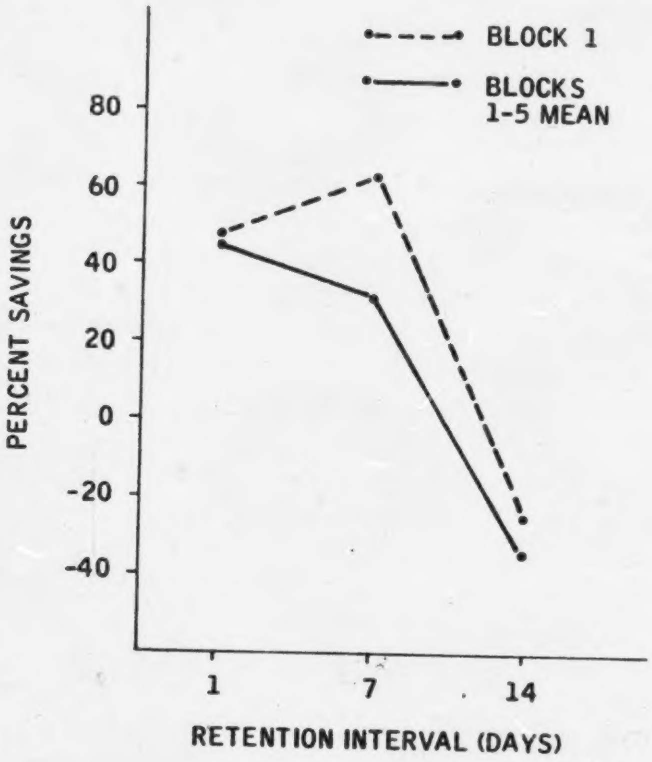


Figure 19. First Block and Average Savings Based on Response Accuracy Scores, as a Function of Retention Interval

64

OVERALL SAVINGS VERSUS FIRST TRIAL SAVINGS

Figures 18 and 19 both show that savings tend to be higher on initial trials than on later trials. This difference is greater for accuracy data than for time data. Further, something akin to a reminiscence effect is obtained on the initial trials for the accuracy data. That is, accuracy on the initial trials is somewhat better after a seven day retention interval than after only a one day interval. Otherwise, the classical retention effect is noted, viz, retention decreases as the interval increases. In both cases, it is significant to note that there are no savings (or worse) with a 14-day retention interval.

These data have basic implications for training procedural tasks in the Navy. It appears that if transfer takes place after more than about a week, training effectiveness might be seriously impaired. Moreover, if the operational task does not provide immediate knowledge of results, errors will tend to perseverate. Thus, in summary, the effects of a retention interval on system effectiveness, under conditions where immediate knowledge of results is not available to the performer, would appear to be a differential function of two factors: (1) the criticality of response speed versus response accuracy to successful performance on the task, and (2) the duration of the retention interval. The present data suggest that the most deleterious effects on retention are manifested on response accuracy after a 14-day retention interval (where performance was degraded close to 35 percent). On the other hand, all other combinations of retention interval and performance measure result in at least as good performance after the retention interval as in the initial learning session. Thus, if system effectiveness depends on accuracy of performance, then retention intervals longer than seven days might lead to significant performance decrements.

SECTION VII

SUMMARY AND RECOMMENDATIONS

Phase II of the L, R, and T program involved four studies. Three of these were laboratory investigations and the fourth was a survey of expert opinion, followed by the synthesis of an approach to studying task difficulty. An evaluation of these activities must be made from the point of view of setting the stage for a long-term experimental program on L, R, and T at NAVTRADEVCEN.

The five experiments conducted during these two initial years of the program concerned themselves solely with the learning, retention, and transfer of a procedural skill. Procedural skills were chosen because they are involved in Naval jobs of all descriptions. It is evident that we have only scratched the surface of knowledge about this category of behavior.

While it appears that verbal learning and verbal behavior underlie many aspects of procedural skills, one cannot account for or predict behavior in the training and transfer of procedural skills, by solely appealing to the list-learning literature. Earlier in this report, a discussion of some of the similarities and differences between procedural and verbal learning was presented. Significantly more research is required in this area before the design of more effective training systems can be accomplished in a scientific manner.

Four variables have been investigated in these five experiments. One of these (imagery) is a possible technique for improving trainees' understanding of basic system structure, and is thus a variable primarily associated with learning. This is not to say, however, that the effects of imagery in training may not be present in retention and transfer. A second variable (fidelity of simulation) refers to similarities and differences between the training and transfer task. Fidelity of simulation can be considered a basic variable relating to transfer. The third variable (retention interval) relates to both learning and transfer in that it influences the extent to which skills acquired in learning sessions are manifested in operational situations. A fourth variable investigated in this series of experiments (individual differences) is related to learning, retention, and transfer by virtue of the fact that the influence of variables associated with all three of these factors are bound up with characteristics of the learner.

IMAGERY

The finding of Phase I, regarding imagery, seemed very promising. Experiment II suggested that the presentation of imagery during training tends to facilitate comprehension of the task. Evidence of this was based on a difference in response latency which favored the imagery group.

Experiments III and IV clearly failed to support the initial findings. In both of these studies, the imagery condition was associated with approximately the same performance levels as were obtained in Experiment II. However, the control groups of the later investigations performed significantly better than the control group of Experiment II.

Imagery has been employed experimentally in verbal learning studies (e. g., Senter and Hauder 1968; Paivio, 1969) to improve acquisition and retention. The findings of such experiments have been very promising. Therefore, it would appear reasonable not to completely dismiss imagery in the procedural skills context.

Several hypotheses about the conditions under which imagery may be effective should be considered further. The first notion regards a possible interaction between the effectiveness of imagery and the abilities of procedural task trainees. This hypothesis was tested in Experiment IV, but due to the lack of spread in the distribution of verbal abilities scores, the data are inconclusive. A similar investigation should be conducted with a wider range of verbal abilities, and possibly, with a battery of tests to find which predict performance on the experimental task given various training methods.

The second hypothesis to explain the inconsistency in findings, regarding imagery, was suggested earlier (see Experiments III and IV). That is, the effectiveness of imagery may interact with the degree of comprehension of the operating instructions. Thus, if subjects understand the task fully without imagery, the introduction of imagery sequences may add little to performances. A test of this hypothesis would require an experiment in which the "goodness" of instruction was varied under conditions of imagery - no imagery.

Finally, the particular way in which imagery was designed may not be optimal. The cartoon sequences involved only about seven minutes of a 90-minute instructional period. Possibly this presentation was not sufficient to provide a significant effect. Alternate ways of providing imagery to trainees to make the task more complete should be considered.

FIDELITY OF SIMULATION

The manipulation of levels of similarity in stimulus and response components between the training and transfer task produced results with important implications to the design of training systems.

Experiments II and V provided data consistent with the notion that transfer depends more on similarity of basic system structure than on stimulus and response similarity. These results were predicted on the basis of Wittrock's (1968) idea of mediated generalization being the basis of transfer. In the present task, subjects recognized that the operational task was not essentially different from the training task and were thus able to reconcile specific differences. Their ability to reconcile differences is very much the same as what psychologists mean by generalization; that is, one stimulus (or response) tending to become equivalent to an earlier stimulus (or response).

The ability to achieve satisfactory training effectiveness under conditions of low fidelity can mean considerable cost savings to the Navy. It does not appear to be necessary to tacitly assume that training devices must be "carbon copies" of their operational counterparts. Instead, designers need to analyze the critical task features of the criterion system to determine how these should be represented in training. If it can reasonably be determined that trainees can recognize correspondences between training and transfer tasks with less than perfect fidelity, considerable cost savings may be possible.

This research by no means allows one to set specifications for fidelity levels. Our findings do suggest, however, that trainees may be able to tolerate rather high degrees of dissimilarity. Further research is needed to determine just how far one may go with this principle.

TASK DIFFICULTY

The decision to survey opinion regarding factors associated with the production of task difficulty was based on the fact that this variable is of central importance in the L, R and T program. As a result of the analysis of the problem, it became obvious that task difficulty is not a unitary variable. Four factors associated with variations in difficulty were identified.

An approach to manipulating these factors was outlined earlier in this report. Basically, we recommend adjusting levels of these factors according to constraints imposed under operational circumstances. For example, if performance of a system has been demonstrated to be unsatisfactory, an analysis of specific reasons for low human performance levels would be required. Improved effectiveness might be achieved by any of the following techniques:

- Improved training methods
- Higher initial skill levels of operational personnel
- Redesign of the task; e.g., greater automation
- Reduced criteria for adequate system performance

Adjustment of all of these factors is not always possible. In situations where one or more of the factors are unalterable, experiments would involve manipulation of the other(s). Such an investigation might involve various methods of simplifying the task or attempting different methods of training. In any case, the nature of the experiment would be dictated by circumstances surrounding the Naval job in question.

The four studies described in this report directly follow the planning and research of the first phase of L, R, and T. It must be recognized that the work accomplished in this two-year effort can be only the beginning of a much larger research program.

The research accomplished thus far has dealt with only one category of behavior. Many of the variables studied are relevant to other taxonomic categories as well. For example, the retention findings for procedural skills may be different for pattern recognition or decision-making skills.

An attempt at deriving a priority list of variables was undertaken in Phase I. Four of these variables were studied in the first two years, viz, imagery, fidelity of simulation, retention interval, and individual differences. The findings associated with fidelity of simulation appear to be the most conclusive of the four. However, even here it is evident that much more research is required before application to the design of training systems can begin to pay off in terms of cost effectiveness.

It is the opinion of the authors that a continuing need will exist to continually establish priorities for the research yet to be conducted. Because economic considerations limit the amount of research conducted, one must continually concentrate on those areas with the greatest promise. It seems safe to say that fidelity of simulation is an area which can be fruitfully investigated. If a decision must be made about limiting research in any one direction, it would appear that further studies of the relationship between cost effectiveness of training and the fidelity of that training would be most worthwhile.

REFERENCES

Ammons, R.B., and Ammons, C.H. The Quick Test (QT): Provisional manual. Psychological Reports, 1962, 11, 111-116 (Monograph Supplement I-VII).

Anderson, J A., Equivalence of meaning among statements presented through various media. AV Communication Review, 1966, 14, 499-505.

Battig, W.F., Facilitation and interference. In E.A. Bilodeau (Ed.), Acquisition of Skill, New York: Academic, 1966.

Battig, W.F., In: Dixon, T.R. and Horton, D.L. (Ed.), Verbal learning and general behavior theory. Englewood Cliff, N.W.: Prentice Hall, 1968.

Bernstein, B.R. and Gonzalez, B.K., Learning, retention, and transfer, Contract N61339-68-C-0216, Honeywell Inc., Research Dept., St. Paul, Minn. September, 1969.

Boltwood, C.E. and Blick, K.A., The delineation and application of three mnemonic techniques. Psychonomic Science, 1970, 20, 339-341.

Bourisseau, W., Davis O.L. Jr., and Yamamoto, K., Sense-impression responses to differing pictorial and verbal stimuli. AV Communication Review, 1965.

Bower, G.H., Organizational factors in memory. Cognitive Psychology, 1970, 1, 18-46.

Bower, G.H., Mental imagery and associative learning. In L. Gregg (Ed.), Cognition in learning and memory. New York: Wiley, in press.

Briggs. G.E. and Naylor, J.C., Team versus individual training, training task fidelity, and task organization effects on transfer performance by three-man teams. Journal of Applied Psychology, 1965, 49, 387-392.

Campbell, J.P., Personnel training and development. Annual Review of Psychology, Vol. 22 (in press).

Conrad, R., Acoustic confusion in immediate memory. British Journal of Psychology, 1964, Vol. 55, 75-84.

Davidson, R.E., and Adams, J.F., Verbal and imagery processes in children's paired-associate learning. Unpublished manuscript. Madison: University of Wisconsin, Undated.

Deese, J., On the structure of associative meaning. Psychology Review, 1962, 69, 161-175.

Demaree, R.G., Norman, D.A., and Matheny, W.G., An experimental program for relating transfer of training to pilot performance and degree of simulation. NAVTRADEVCEN 1388-1, Contract N61339-1338, Life Sciences, Inc. June 1965.

Deno, S.L., Effects of words and pictures as stimuli in learning language equivalents. Journal of Educational Psychology, 1968, 59, 202-206.

Deno, S.L., Johnson, P.E. and Jenkins, J.R., Associative similarity of words and pictures. AV Communication Review, 1968, 16, 280-286.

Ellis, N.C., Lowes, A.L., Matheny, W.G., and Norman, D.A., Pilot performance, transfer of training and degree of simulation: III. Performance of non-jet experienced pilots versus simulation fidelity. NAVTRADEVCEN 67-C-0034-1, Contract N61339-67-C-0034, Inc. Aug. 1968.

Fleishman, E.A., Individual differences and motor learning. In R.M. Gagne (Ed.), Learning and individual differences. Columbus, Ohio: Merrill, 1967.

Folley, J.F. Jr., and Munger, S.J.,,Training performance aids - A review of the literature on the design of informational job performance aids. USAF ASD TR 61-549. Oct., 1961, 1-43.

Gagne, R.M., Simulators, In. R. Glaser (Ed.), Training research and education. University of Pittsburgh Press, 1962.

Grimsley, D.L., Acquisition retention and retraining: Effects of high and low fidelity in training devices. The George Washington University, Human Resources Research Office. Technical Report 69-1, 1969.

Hammerton, M., and Tickner, A.H., Visual factors affecting transfer of training from a simulated to a real control situation. Journal of Applied Science, 1967, 51, 46-49.

Jenkins, J.R., Neale, D.C., and Deno, S.L., Differential memory for picture and word stimuli. Journal of Educational Psychology, 1967, 58, 303-307.

Kale, S.V., Grosslight, J.H., and McIntyre, C.J., Exploratory studies in the use of pictures and sound for teaching foreign language vocabulary. Office of Naval Research, 1955.

Koehnke, K.R., The roles of pictures and readability in comprehension of the main idea of a paragraph. Paper presented at the annual Meetings of the American Education Research Association, Chicago, February, 1968.

Lieberman, L.R , and Culpepper, J.T , Words versus objects: Comparison of free verbal recall. *Psychological Reports*, 1965, *17*, 983-988.

Mahler, W.R., and Bennett, G.K., *Psychological studies of advanced naval training: Evaluation of operational flight trainers*. Port Washington, N.Y.: *USN Special Devices Center*. Technical Report SDC 999-1-1, 1950.

Milgram, N.A., Verbal context versus visual compound in paired-associate learning by children. *Journal of Experimental Child Psychology*, 1967.

Montague, W.E., Adams, J.A., and Kiess, H.O., Forgetting and natural language mediation. *Journal of Experimental Psychology*, 1966, *72*, 829-833.

Montague, W.E., and Wearing, A. J., The complexity of natural language mediators and its relation to paired-associate learning. *Psychonomic Science*, 1967, *7*, 135-136.

Newton, J.M., *Training effectiveness as a function of simulator complexity*, *Technical Report NAVTRADEVCEN 458-1, Training Device Center*, Port Washington, N.Y., September 1959.

Osgood, C.E., The similarity paradox in human learning: A resolution. *Psychological Review*, 1949, 56, 132-143.

Paivio, A., Learning of adjective-noun paired associates as a function of adjective-noun word order and noun abstractness. *Canadian Journal of Psychology*, 1963, *17*, 370-370.

Paivio, A., Mental imagery in associative learning and memory. *Psychological Review*, 1969, *76*, 241-263.

Paivio, A. On the functional significance of imagery. *Psychological Bulletin*, 1970, *73*, 385-392.

Paivio, A., and Yarmey, A.D., Pictures versus words as stimuli and responses in paired-associate learning. *Psychonomic Science*, 1966, *5*, 235-236.

Paivio, A., Rogers, T.B., and Smythe, P.C., Why are pictures easier to recall than words? *Psychonomic Science*, 1968, *11*, 137-138.

Paivio, A., Yuille, J C., and Madigan, S., Concreteness, imagery and meaningfulness values for 925 nouns. *Journal of Experimental Psychology monograph Supplement*, Part 2, 1968, *76, 1-25*.

Paivio, A., Yuille, J. C., and Smythe, P. C. Stimulus and response abstractness, imagery, and meaningfulness, and reported mediators in paired-associate learning. *Canadian Journal of Psychology*, 1966, *20*, 362-377.

Reese, H.W., Implications of mnemonics research for cognitive theory. Paper presented at the Southeastern Conference on Research in Child Development, Athens, Georgia, April 1970.

Reese, H.W., Imagery and contextual meaning. Psychological Bulletin, 1970, 73, 393-403.

Rohwer, W.D., Jr. Constraint, syntax and meaning in paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 1966, 5, trl-547.

Rohwer, W.D., Jr., Images and pictures in children's learning. Psychological Bulletin, 1970, 73, 385-392.

Samuels, S.J. Effects of pictures on learning to read, comprehension and attitudes. Review of Educational Research, 1970, 40, 397-408.

Senter, R.J., and Hauder, G., An experimental study of a mnemonic system. Psychonomic Science, 1968, 10, 289-290.

Shepard, R.N. and Chipman S., Second-order isomorphism of internal representations: shapes of states. Cognitive Psychology, 1970, 1, 1-17.

Silver, C.A., Jones, J.M. and Landis, D., Analysis of radar training requirements. NAVTRADEVCEN 1345-1, Contract N61339-1345, Franklin Institute Research Laboratories, Aug. 1965, 37pp. AD 472414.

Tulving, E., Theoretical issues in free recall. In Dixon, T.R. and Horton, D.L. (Eds.), Verbal behavior and general behavior theory. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

Wilcoxon, H.C., Davy, E., and Webster, J.C., Evaluation of the SNJ operational flight trainer. Port Washington, N.W.: USN Special Devices Center, Report 999-2-1, 1954.

Willis, M.P., and Peterson, R.O. Deriving training device implications from learning theory principles: Vol. II: Methodology. Technical Report: NAVTRADEVCEN 784-2, Port Washington, N.Y. 1961, II, 1-32.

Wittrock, M.C. Three conceptual approaches to research. In R.M. Gagne and W. J. Gephart (Eds.) Learning research and school subjects. Itasco, Illinois: Peacock. 1968.

Woodworth, R.S. and Schlosberg, H. Experimental psychology. New York: Henry Holt, 1954.

Yates, F.A., The art of memory, London: Routledge and Kegan Paul, 1966.

APPENDIX A

SUMMARY OF INTERVIEW ON TASK DIFFICULTY

Interviewee	Factors Affecting Difficulty
H. Ammerman	• Number of decision options or branches in a procedure (i. e., troubleshooting versus fixed procedure).
	• (See additional items under E. Miller)
G. Briggs	• Number of different things a person must do at any single time.
	• Within any single dimension, the amount and kind of processing required.
	• Task organization, i. e., degree of nonindependence with dimensions.
G. Jeantheau	• Relative geometry of units, i. e., the fact, rate and amount of change.
	• Physical environment as it influences sensors.
	• System load
	- number of units to be dealt with
	- communication and coordination requirements
	• Operational mission
	- threat, urgency
	• Contingencies, emergencies, malfunctions.
R. Mackie	• Signal to noise ratio
	• Multiple stimuli representing same object
	• Equipment alignment
	• Environment

T. Mara	• Level of ambiguity • Amount of information available • Amount of mental storage available
G. Matheny	• Overloading - time stress - speed and accuracy requirements • Psychological stress • Adaptation requirements, i. e., adapting to a variety of situations and environmental conditions
E. Miller	• Procedural memory (time from interactions to actual performance) • Time stress (speed requirements) • Visual acuity (in raw video displays) • Task distractor stress (doing two things at once) • Environmental distractors (danger, noise, fear) • Intersystem coordination (timing and phasing) • Concept complexity (coordination and understanding what others are doing with your job) • General complexity (length of process) • Physical strength stress
L. Schrenk	• Number of information input channels used • Rate of information input • Uncertainty or unpredictability of input task • Complexity of rules for information processing • Complexity - amount of memory required

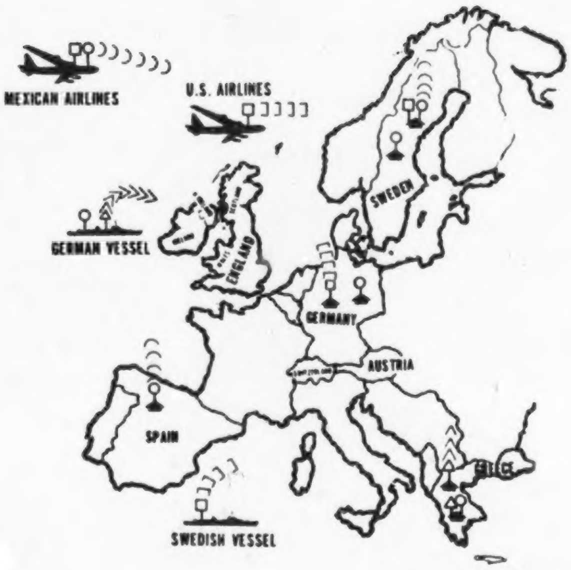
A. Siegel

- Environmental factors - number of targets, target speed, etc.
- Crew integration requirements
- Intellectual loading

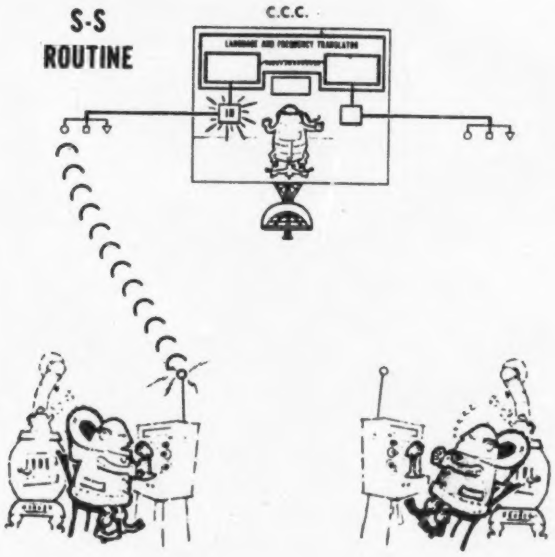
B. Yaeger

- Stimulus difficulty discrimination
- Cognitive difficulty
  - procedure complexity
  - procedure selection
- Response difficulty
  - timing
  - accuracy
  - rhythm

APPENDIX B
DIRECT IMAGERY (L, R, AND T)

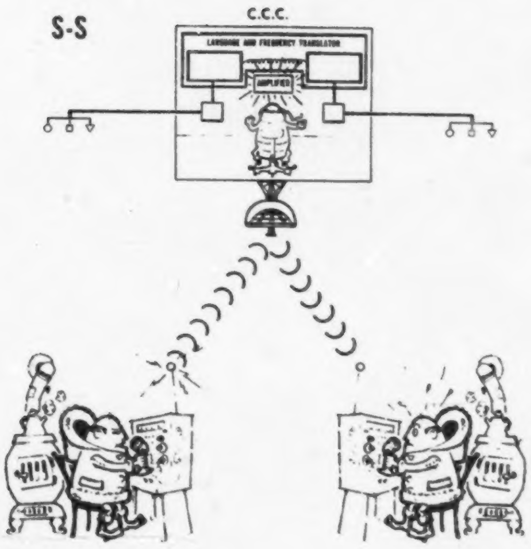

MEXICAN AIRLINES
U.S. AIRLINES
GERMAN VESSEL
ENGLAND
SWEDEN
GERMANY
AUSTRIA
SPAIN
GREECE
SWEDISH VESSEL

First, there is a map of Western Europe, parts of the Mediterranean, and the Atlantic Ocean. It shows the countries which speak the five languages included in this study: English, German, Spanish, Greek and Swedish. Also evident are some ships and aircraft which communicate in these languages. The land-based stations, the airborne ones, and those on shipboard have one, two, or three types of antennae -- of the three types of antennae, each emits its own unique type of signal.

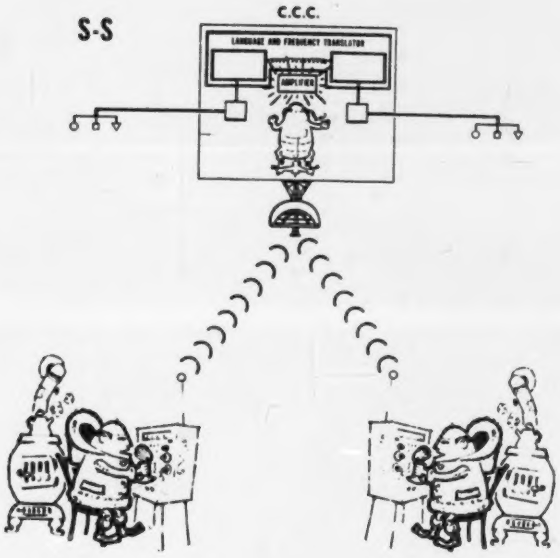

S-S
ROUTINE
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR

The S-S routine can be described in 3 steps:

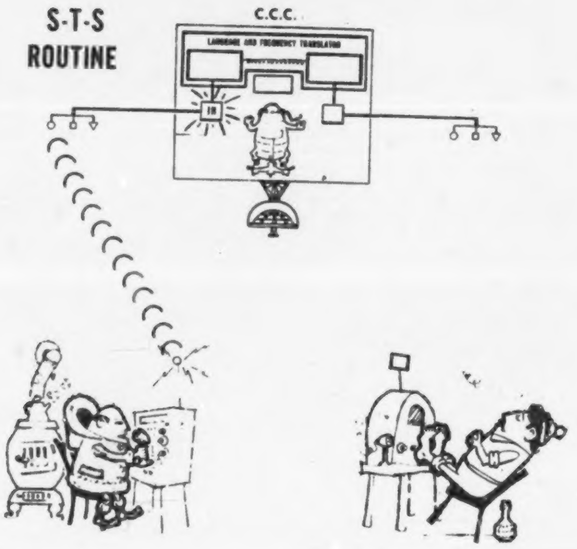
First, Sven Peterson, the operator of a Swedish Arctic outpost has a message for Christen Olstad, manning the communication facility in the Swedish port of Holmstad. Peterson, the sender, flashes his communication request to the Communications Control Console.


S-S
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR
AMPLIFIER

In the second scene, Peterson's request is received by the Control Console Operator. The Operator quickly determines that both Peterson and Olstad, the intended receiver, speak Swedish, and that both operate HF radio equipment. So, the console Operator relays the message directly from Peterson to Olstad on an available HF channel.


S-S
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR
AMPLIFIER

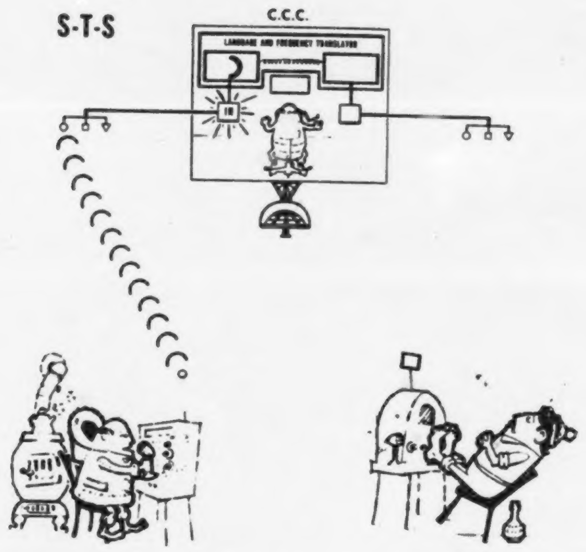
Finally, the connection between Peterson and Olstad is made, and they are in direct communication with one another, with the radio signals being amplified through the Communication Control Console.



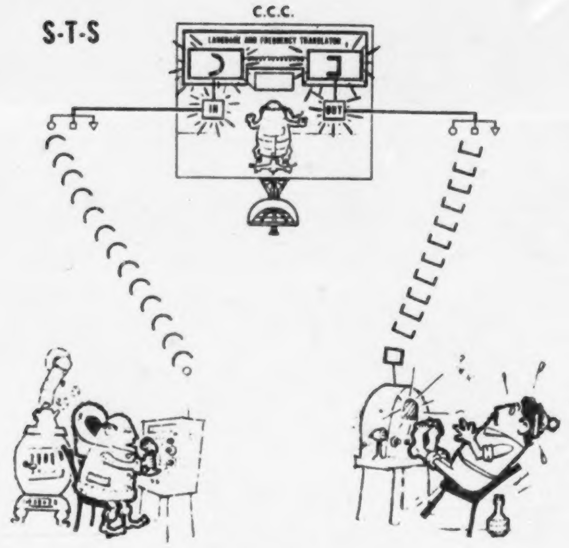
The S-T-S routine can be described in the following 4 scenes:

Joris Swenson in Stockholm wants to communicate with Tully Spyropolous, in Greece. The Communication Control Console Operator receives the message request from Swenson.

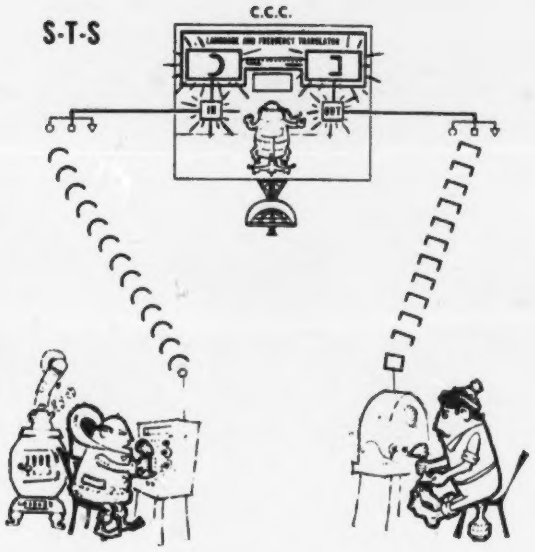
The Console Operator checks the language and frequency capabilities of the Sender and the intended Receiver. He ascertains that Swenson speaks Swedish and has HF equipment. Spyropolous, on the other hand, speaks Greek and has only VHF radio equipment. The Console Operator, therefore, has to process the message STS, through the Language and Frequency Translator.


S-T-S
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR
HF

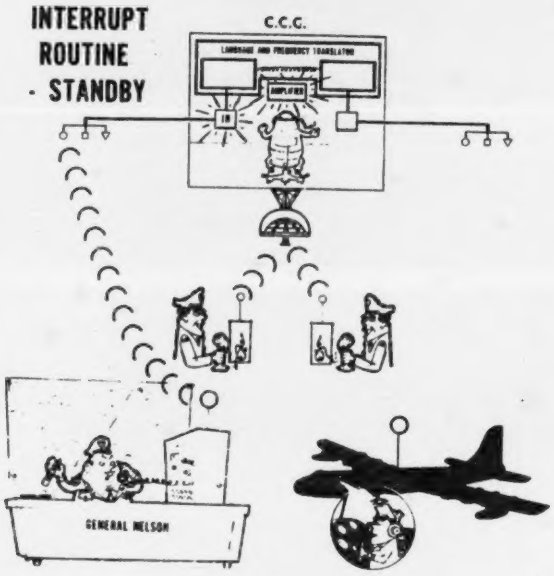
In the next scene, the Control Console Operator brings the message from Swenson into the console, on HF frequency. The message is then fed automatically into the Language and Frequency Translator.


S-T-S
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR
IN
OUT

Next, the Operator redirects the message to the Receiver, Spyropolous, on an open VHF channel.


S-T-S
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR
IN
OUT

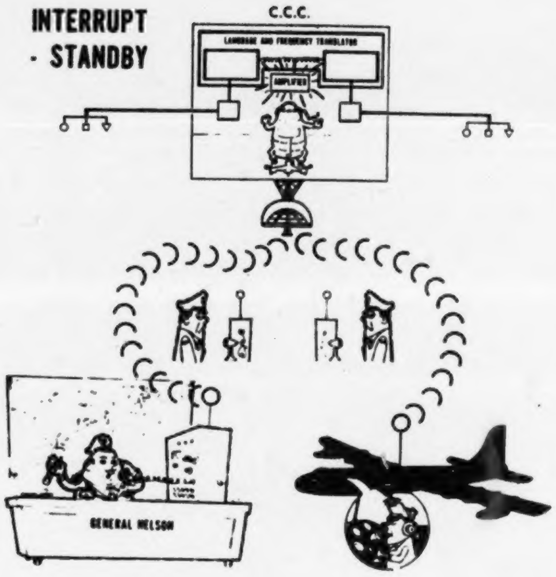
Finally, at this point, Swenson, with HF equipment, and Spyropolous, with VHF equipment, are communicating freely with one another through the Translator.


INTERRUPT
ROUTINE
- STANDBY
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR
AMPLIFIER
IN
GENERAL NELSON

Next we have the INTERRUPT ROUTINE, first for Interrupt-Standby.

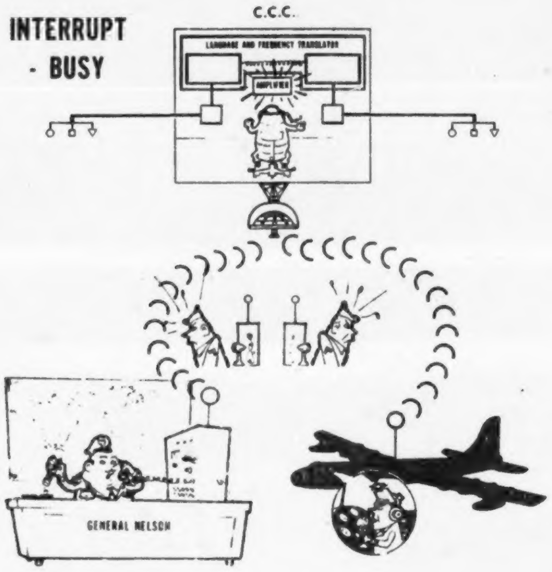
This situation shows General Nelson, who has an extremely important message for Colonel Maxwell who is aboard an aircraft. The Console Operator sees that the only communication channel usable by the Senior Officers is currently being used by two lieutenants conveying a fairly important message.

Because of the urgent nature of the General's call, the Console Operator decides that he must interrupt the lieutenants.


INTERRUPT
- STANDBY
C.C.C.
AMPLIFIER
GENERAL NELSON

In the next scene, the Console Operator flashes the standby signal to the lieutenants. The lieutenants, upon receiving the standby signal, temporarily discontinue their communication.

The channel is thus freed and the Operator uses it to complete the patch between General Nelson and Colonel Maxwell.

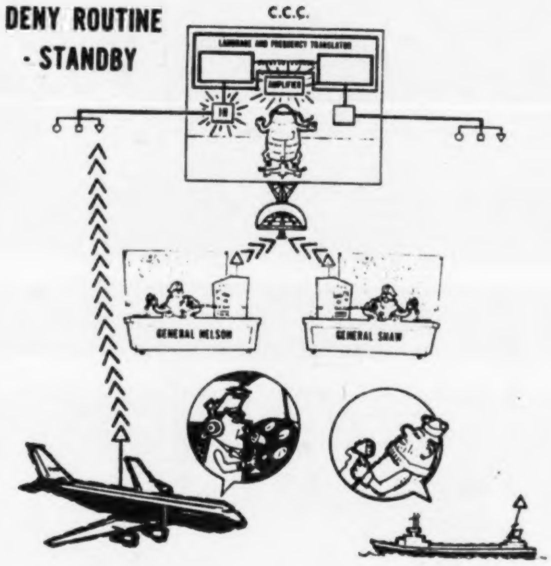

INTERRUPT
- BUSY
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR
AMPLIFIER
GENERAL NELSON

Next, we have INTERRUPT - BUSY

The next time General Nelson had an important communique for Colonel Maxwell, the available channel is being used by two PFCs conveying routine information.

The Console Operator flashes a BUSY signal to the PFCs indicating that they are to be cut off.

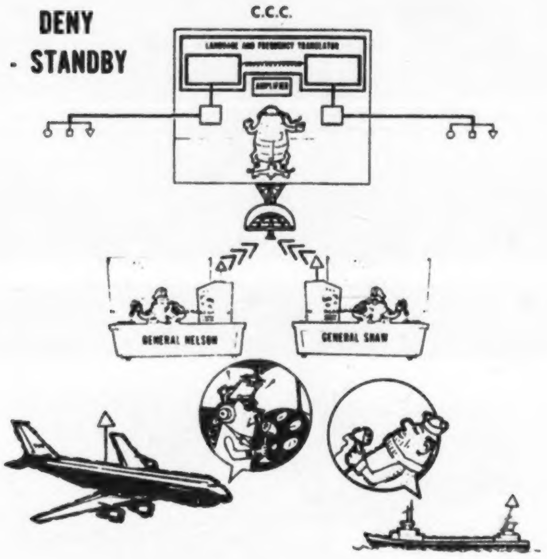
Therefore, the PFCs will have to re-initiate their communication later.


DENY ROUTINE
- STANDBY
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR
AMPLIFIER
10
GENERAL NELSON
GENERAL SHAW

Next is DENY - STANDBY

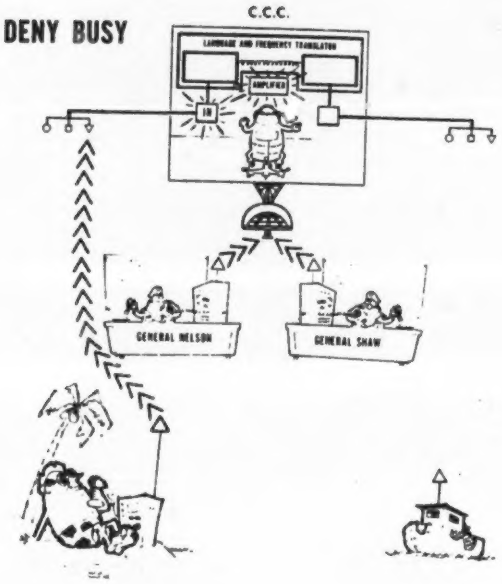
Captain Breckenridge, piloting a KC-135, has a very important message to deliver to the USS Hoover, which is with the 7th fleet, off Gibraltar.

The Console Operator notes that the only channel the ship can receive messages on is currently being used by two senior officers whose communication is just as important as that of Captain Breckenridge.


DENY
- STANDBY
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR
AMPLIFIER
GENERAL NELSON
GENERAL SHAW

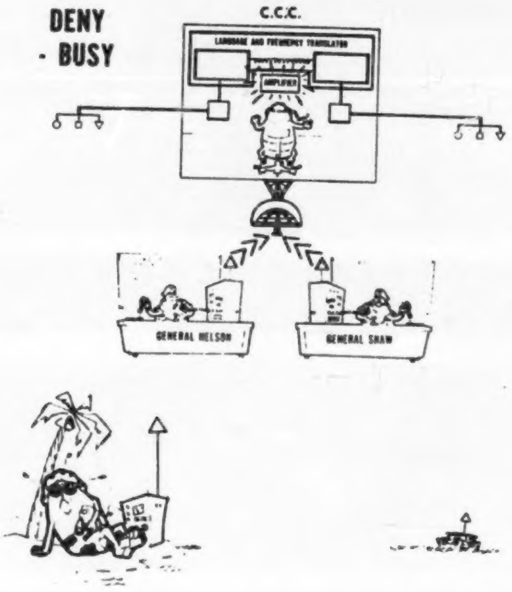
Therefore, the next step would be for the Console Operator to deny Captain Breckenridge's request to communicate.

Because Captain Breckenridge had a high priority request, he is placed on STANDBY.


DENY BUSY
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR
AMPLIFIER
IN
GENERAL NELSON
GENERAL SHAW

The last situation is DENY - BUSY

Commander Conklin, off duty, on the Riviera, remembers a few details he wants his aide, aboard a small service boat, to clear up for him. The Commander, therefore, requests a patch between himself and his aide. The Console Operator sees that all available communication channels are currently in use by other parties.


DENY
- BUSY
C.C.C.
LANGUAGE AND FREQUENCY TRANSLATOR
AMPLIFIER
GENERAL NELSON
GENERAL SHAW

So, since Conklin's message is clearly of low priority, the operator flashes him a busy signal and he will have to re-request a patch at some other time.

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)
Honeywell Inc.
Systems and Research Division, Research Department
St. Paul, Minnesota 55113

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP
NA

3. REPORT TITLE
LEARNING, RETENTION AND TRANSFER IN MILITARY TRAINING

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Final Technical Report, 23 June 1969 through 23 June 1970

5. AUTHOR(S) (First name, middle initial, last name)
Bernard R. Bernstein, Barbara K. Gonzalez

6. REPORT DATE
September 1971

7a. TOTAL NO. OF PAGES
82

7b. NO. OF REFS
54

8a. CONTRACT OR GRANT NO.
N61339-69-C-0253

b. PROJECT NO.
7846-2

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)
12174-FR

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
NAVTRADEVCEN 69-C-0253-1

10. DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY
Naval Training Device Center
Orlando, Florida

13. ABSTRACT

This report summarizes Phase I and reports on Phase II of the L, R, and T program. The purposes of Phase II were to analyze the factors contributing to task difficulty in Naval jobs, and to conduct a further exploration of the influence of imagery and fidelity of simulation on learning, retention, and transfer. A survey of expert opinion regarding task difficulty plus three experimental investigations were conducted during this one-year effort.

The specific purpose of the Task Difficulty Survey was to define those factors which significantly influence task difficulty in the operational Navy setting. Once identified, these factors can be used to realistically manipulate the difficulty of tasks in laboratory investigations. Telephone interviews were conducted with 10 scientists, representing both military and civilian research organizations, to obtain the required information.

The results of this survey were inconclusive. Further, a majority of the difficulty factors mentioned by interviewees were the same ones which have been previously studied in the laboratory. Consequently, the problem became one of suggesting an approach for manipulating such variables. An overall plan is suggested to guide subsequent investigations in which task difficulty is either controlled or maniputed.

Experiments III, IV and V of the current series were conducted during Phase II. Experiments III and IV dealt with imagery and Experiment V with fidelity of simulation.

DD FORM 1 NOV 65 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Learning						
Retention						
Transfer of training						
Training research						
Fidelity of simulation						
Imagery						
Procedure-following task						

Experiment III addressed two questions:

1. What is the relative training effectiveness of two forms of imagery; i.e., (a) when new material is related to more familiar material in an analogous fashion and (b) when new material is represented directly?

2. What are the relative contributions of images and text in the presentation of imagery?

This experiment was designed to validate the facilitating influence of imagery found in Experiment II of the series. The findings of Experiment III failed to support the facilitation found in Experiment II. These results are discussed in light of the subject samples and experimental procedures used.

Experiment IV was similar to III, with the exception that subjects were trained with a paper and pencil rather than a hardware simulation of the task. The purposes of this study were to resolve discrepancies in the results between Experiments II and III, and to test the hypothesis that the effectiveness of imagery interacts with verbal skills. The findings were consistent with Experiment III providing no evidence that imagery improves performance. This finding was consistent regardless of verbal skill level. The hypothesis that imagery should be more effective for subjects with lower verbal skills was thus not supported. Those results are discussed in relation to the restricted range of verbal skills of the subjects used.

Experiment V investigated the influence of training task fidelity of simulation on transfer performance. Subjects were trained using one of three levels of fidelity. The data indicated that variations in stimulus and response fidelity had little effect on response time or accuracy. The data are interpreted as supporting the findings of Experiment II; i.e., that relatively high levels of training effectiveness can be obtained for procedural tasks in the absence of high physical fidelity between training and transfer tasks.

In a final section the work on this program is summarized and recommendations made for the direction of future research.

Naval Training Device Center, Orlando, Fla.

TR NAVTRADEVCEN 69-C-0253-1 UNCLASSIFIED
AD________

LEARNING, RETENTION, AND TRANSFER IN MILITARY TRAINING. 1971, 82 p., 19 illus., 14 tables, 54 refs, 2 append.

Investigations into the role of imagery and fidelity of simulation on learning retention and transfer were conducted. These experiments were part of a long range program leading to improved training methods for the Navy. Imagery did not have a consistently beneficial effect on transfer performance. Relatively high levels of training effectiveness resulted for procedural tasks in the absence of high physical fidelity between training and transfer conditions.

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B. R.
and others
N61339-69-C-0253.

Naval Training Device Center, Orlando, Fla.

TR NAVTRADEVCEN 69-C-0253-1 UNCLASSIFIED
AD________

LEARNING, RETENTION, AND TRANSFER IN MILITARY TRAINING. 1971, 82 p., 19 illus., 14 tables, 54 refs, 2 append.

Investigations into the role of imagery and fidelity of simulation on learning retention and transfer were conducted. These experiments were part of a long range program leading to improved training methods for the Navy. Imagery did not have a consistently beneficial effect on transfer performance. Relatively high levels of training effectiveness resulted for procedural tasks in the absence of high physical fidelity between training and transfer conditions.

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B. R.
and others
N61339-69-C-0253.

Fla.

LASSIFIED

ER IN
illus.,

d fidelity
ansfer
e part of
d training
ave a
perfor-
effec-
the
training

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B. R.
and others
N61339-69-C-0253.

Naval Training Device Center, Orlando, Fla.

TR NAVTRADEVCEN 69-C-0253-1 UNCLASSIFIED
AD__________

LEARNING, RETENTION, AND TRANSFER IN MILITARY TRAINING. 1971, 82 p., 19 illus., 14 tables, 54 refs, 2 append.

Investigations into the role of imagery and fidelity of simulation on learning retention and transfer were conducted. These experiments were part of a long range program leading to improved training methods for the Navy. Imagery did not have a consistently beneficial effect on transfer performance. Relatively high levels of training effectiveness resulted for procedural tasks in the absence of high physical fidelity between training and transfer conditions.

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B. R.
and others
N61339-69-C-0253.

Fla.

LASSIFIED

ER IN
illus.,

d fidelity
ransfer
e part of
d training
have a
perfor-
g effec-
the
training

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B. R.
and others
N61339-69-C-0253.

Naval Training Device Center, Orlando, Fla.

TR NAVTRADEVCEN 69-C-0253-1 UNCLASSIFIED
AD__________

LEARNING, RETENTION, AND TRANSFER IN MILITARY TRAINING. 1971, 82 p., 19 illus., 14 tables, 54 refs, 2 append.

Investigations into the role of imagery and fidelity of simulation on learning retention and transfer were conducted. These experiments were part of a long range program leading to improved training methods for the Navy. Imagery did not have a consistently beneficial effect on transfer performance. Relatively high levels of training effectiveness resulted for procedural tasks in the absence of high physical fidelity between training and transfer conditions.

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B. R.
and others
N61339-69-C-0253.

Naval Training Device Center, Orlando, Fla.

TR NAVTRADEVCEN 69-C-0253-1 UNCLASSIFIED
AD__________

LEARNING, RETENTION, AND TRANSFER IN MILITARY TRAINING. 1971, 82 p., 19 illus., 14 tables, 54 refs, 2 append.

Investigations into the role of imagery and fidelity of simulation on learning retention and transfer were conducted. These experiments were part of a long range program leading to improved training methods for the Navy. Imagery did not have a consistently beneficial effect on transfer performance. Relatively high levels of training effectiveness resulted for procedural tasks in the absence of high physical fidelity between training and transfer conditions.

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B. R.
and others
N61339-69-C-0253.

Naval Training Device Center, Orlando, Fla.

TR NAVTRADEVCEN 69-C-0253-1 UNCLASSIFIED
AD__________

LEARNING, RETENTION, AND TRANSFER IN MILITARY TRAINING. 1971, 82 p., 19 illus., 14 tables, 54 refs, 2 append.

Investigations into the rôle of imagery and fidelity of simulation on learning retention and transfer were conducted. These experiments were part of a long range program leading to improved training methods for the Navy. Imagery did not have a consistently beneficial effect on transfer performance. Relatively high levels of training effectiveness resulted for procedural tasks in the absence of high physical fidelity between training and transfer conditions.

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B. R.
and others
N61339-69-C-0253.

ndo, Fla.

UNCLASSIFIED
AD____________

ANSFER IN
, 19 illus.,

ry and fidelity
nd transfer
s were part of
roved training
not have a
sfer perfor-
aining effec-
ks in the
ween training

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B.R.
and others
N61339-69-C-0253.

Naval Training Device Center, Orlando, Fla.

TR NAVTRADEVCEN 69-C-0253-1 UNCLASSIFIED
AD____________

LEARNING, RETENTION, AND TRANSFER IN MILITARY TRAINING. 1971, 82 p., 19 illus., 14 tables, 54 refs, 2 append.

Investigations into the role of imagery and fidelity of simulation on learning retention and transfer were conducted. These experiments were part of a long range program leading to improved training methods for the Navy. Imagery did not have a consistently beneficial effect on transfer performance. Relatively high levels of training effectiveness resulted for procedural tasks in the absence of high physical fidelity between training and transfer conditions.

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B.R.
and others
N61339-69-C-0253.

ndo, Fla.

UNCLASSIFIED
AD____________

ANSFER IN
, 19 illus.,

ry and fidelity
nd transfer
s were part of
roved training
not have a
sfer perfor-
aining effec-
ks in the
ween training

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B.R.
and others
N61339-69-C-0253.

Naval Training Device Center, Orlando, Fla.

TR NAVTRADEVCEN 69-C-0253-1 UNCLASSIFIED
AD____________

LEARNING, RETENTION, AND TRANSFER IN MILITARY TRAINING. 1971, 82 p., 19 illus., 14 tables, 54 refs, 2 append.

Investigations into the role of imagery and fidelity of simulation on learning retention and transfer were conducted. These experiments were part of a long range program leading to improved training methods for the Navy. Imagery did not have a consistently beneficial effect on transfer performance. Relatively high levels of training effectiveness resulted for procedural tasks in the absence of high physical fidelity between training and transfer conditions.

DESCRIPTORS

Learning
Retention
Transfer of training
Fidelity of simulation
Imagery
Procedure-following task

Honeywell Inc.
Bernstein, B.R.
and others
N61339-69-C-0253.